

FIG. 1

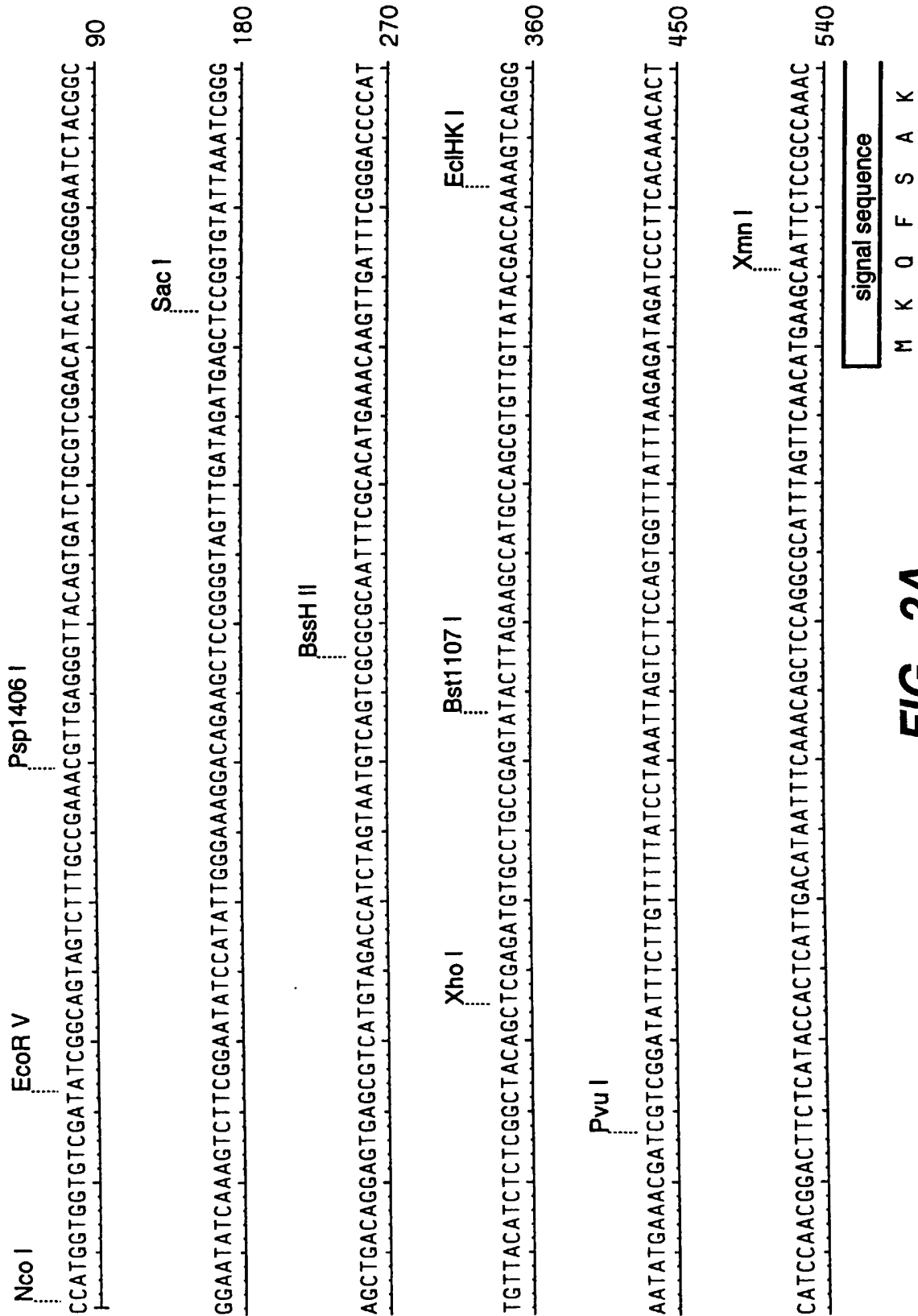


FIG._2A

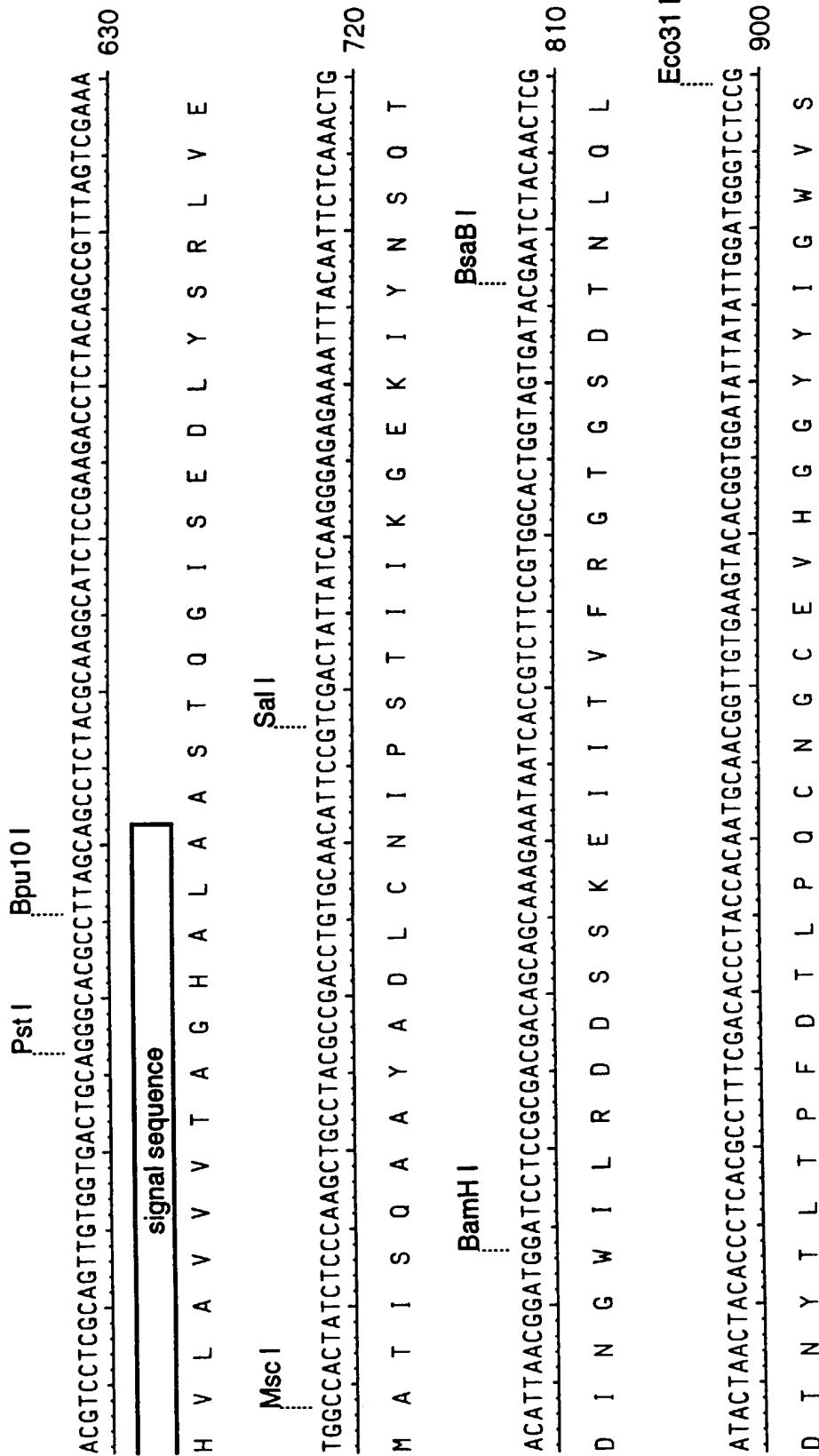


FIG..2B

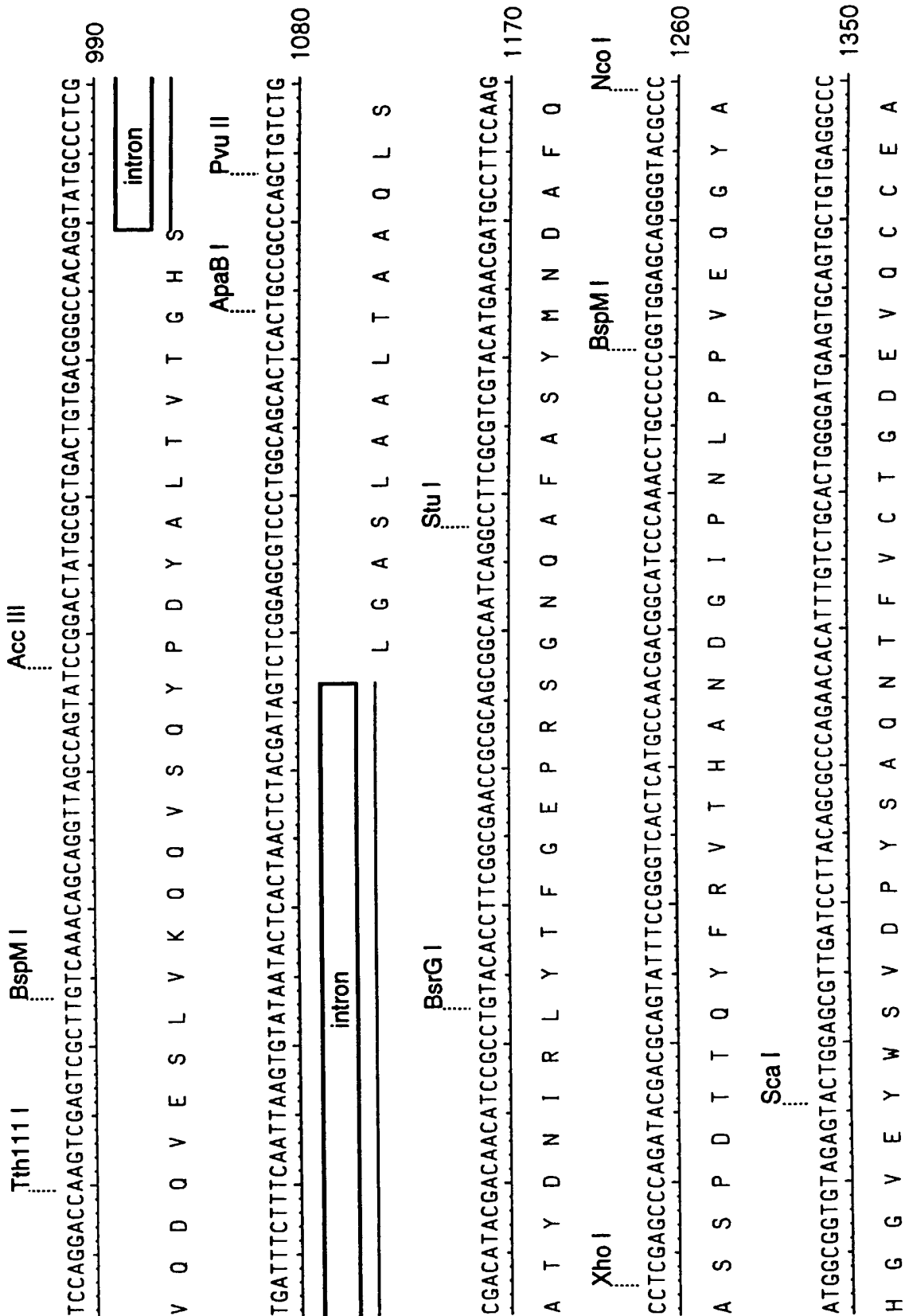


FIG._2C

5 / 154

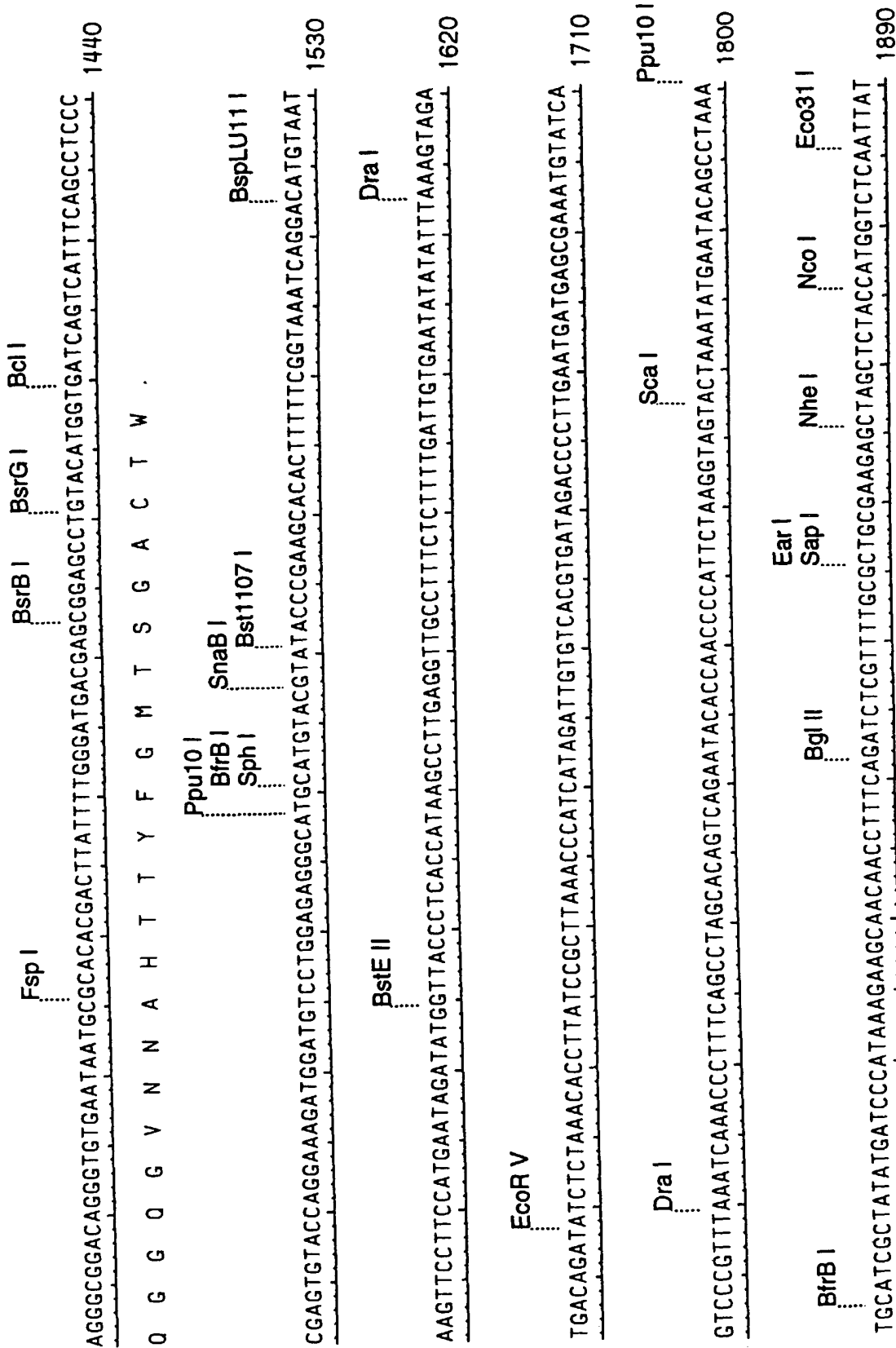


FIG.-2D

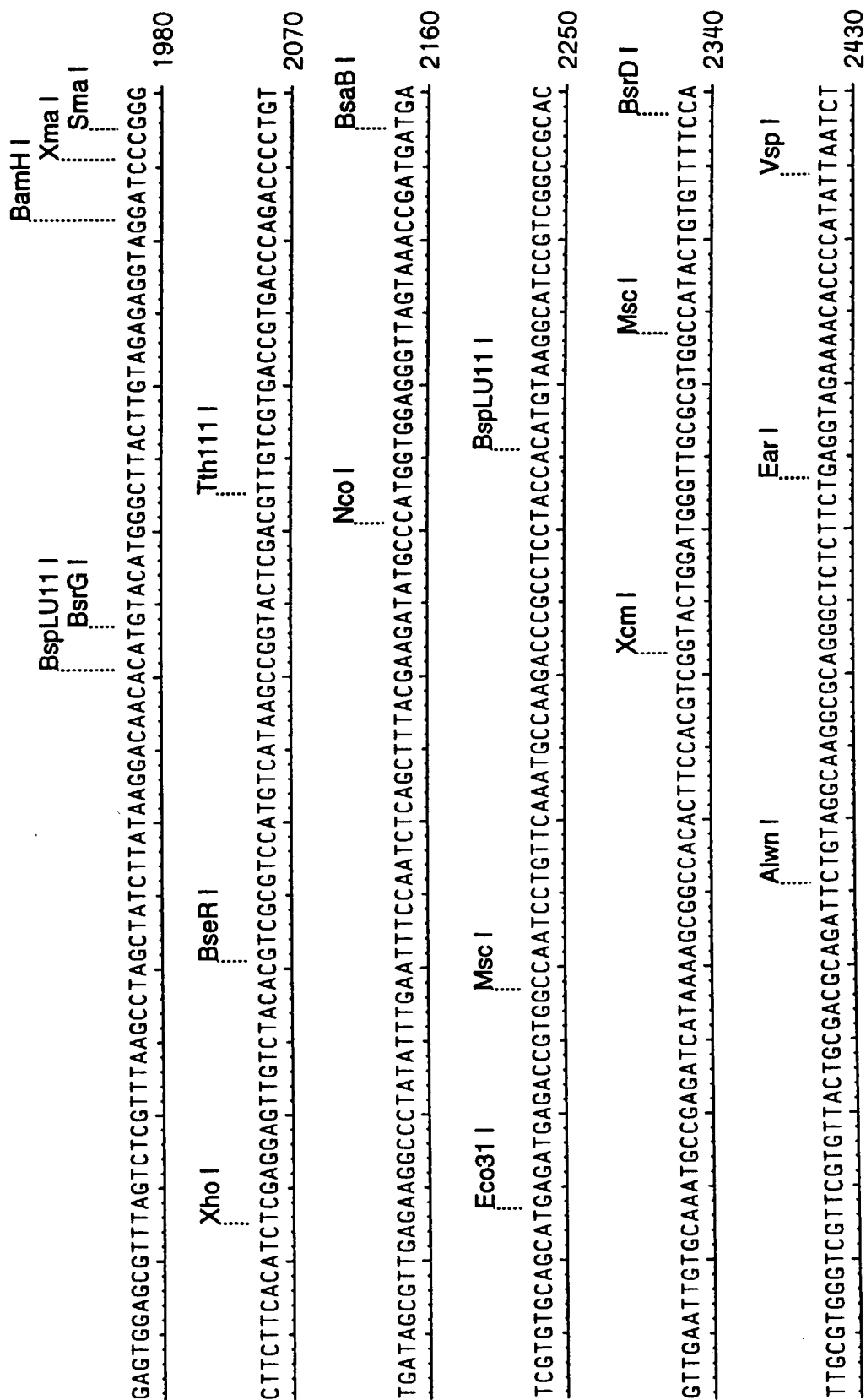


FIG._2E

CCATGGTGGTGCATATCGGCAGTAGTCTTTGCCGAACGTTGAGGGTTACAGTGATCTGCGTCGGACATACATT
CGGGGAATCTACGGCGGAATATCAAAGTCTTCGGAATATCCATATTTGGAAAGGACAGAAAGCTCCGGGTAGTTT
GATAGATGAGCTCCGGTGATTAATCGGGAGCTGACAGGAGTGAGCGTCATGTAGACCATTAGTAATGTCAAGT
CGCGGCCAATTTCGCACATGAACAAGTTGATTTCGGGACCCCATTTGTTACATCTCTCGGCTACAGCTCGAGATG
TGCCCTGCCGAGTACTTAGAAGCCATGCCAGCGTGTGTTTATACGACCAAAAGTCAAGGAATATGAACGATCG
TTCGGATATTTCTTGTGTTTATCCTAAATTAGTCTTCCAGTGGTTTATTTAAGAGATAGATCCCCTTCACAAACACT
CATCCAAACGGACTTCTCATACCACTCATTTGACATAATTTCAAACAGCTCCAGGCGCATTTAGTTCAACATGAAGC
AAATCTCCGCCAAACAGTCTCGCAGTTGTGGTGACTGCAAGGCACGCCTTAGCAGCTCTACGCAAGGCATCT
CCGAAGACCTCTACAGCCGTTTAGTCGAAATGGCCACTATCTCCAAAGCTGCCACGCCGACTGTGCAACATTC
CGTCGACTATTATCAAGGAGAGAAAATTTACAATTCTCAAACCTGACATTAACGGATGGATCCTCCGCGACGACA
GGAGCAAGAAATAATCACCGTCTTCGGTGGCACTGGTAGTATACGAATCTACAACCTCGATATACTACACCC
TTCACGCCCTTTCGACACCTACCAATGCAACGGTGTGAAATACAGGTGGATATTATATTGGATGGGTCTCCG
TCCAGGACCAAGTCGAGTCGCTTGTCAAACAGCAGGTTAGCCAGTATCCGGACTATGCGCTGACTGTGACGGGCC
ACAGGTAATGCCCTCGTGATTCTTTCAAATTAAGTGATAATACTCACTAACTACGATAGTCTCGGAGCGTCCC
TGGCAGCACTACCTGCCGCCAGCTGTCTGCGACATACGACAACTACGCCCTGTACACCTTCGGCGAACCGCGCA
GCGGCAATCAGGCCCTTCGCTCGTACATGAACGATGCCCTTCCAAGCCTCGAGCCAGATACGACGCAGTATTTC
GGGTCACTCATGCCCAACGACGGCATCCCAAACCTGCCCTCGGTGGAGCAGGGGTACGCCCATGGCGGTGTAGAGT
ACTGGAGCGTTGATCCTTACAGGCCCAAGAACACATTTGTCTGCACCTGGGGATGAAGTGCAGTGTGTGAGGCC
AGGGCGGACAGGGTGTGAATAATCGGCACACGACTTATTTGGGATGACGAGCGGAGCCTGTACATGGTGATCAG
TCATTTCAAGCCTCCCGAGTGTAACAGGAAAGATGGATGTCTTGAGAGGGCATGCTACGTATACCCGAAGC
ACACTTTTCGGTAAATCAGGACATGTAATAAGTTCTCTCCATGAATAGATATGGTTACCTCACCATAGCCTT
GAGGTGCGCTTCTCTTTGATTGTGAAATATAATTAAGTAGATGACAGATATCTCTAAACACCTTATCCGCT
TAAACCCATCATAGATTGTGTACGTAGACCCCTTGAAATGATGAGCGAAATGTATCAGTCCCCTTTAAATCA
AACCTTTCAAGCTAGCACAGTCAGAAATACACCAACCCCATTTCTAAGGTAGTACTAAATATGAATACAGCCTAAA
TGCAATCGCTATATGATCCCATAAAGAAGCAACACCTTTCAGATCTCGTTTTCGCTGCGAAGAGCTAGCTCTAC
CATGGTCTCAATTATGAGTGGAGCGTTTAGTCTCGTTTAAAGCCTAGCTATCTTATAAGGACAAACATGTACATG
GGCTACTGTAGAGAGGTAGGATCCCGGGCTTCTTCACATCTCGAGGAGTTGTCTACACGTGCGCTCCATGTCA
TAAGCCGGTACTCGACGTTGTGTCGTGACCGTGACCCAGACCCCTGTTGATAGCGTTGAGAAGGCCCTATATTGAA
TTTCCAACTCAGCTTTACGAAGATATGCCCATGGTGGAGGTTAGTAAACCGATGATGATCGTGTGACGATGA
GATGAGACCGTGGCCAATCCTGTTCAAATGCCAAAGACCCGCCCTCTACCACATGTAAAGCATCCGTGCGCCGCA
GTTGAATTGTGCAAAATGCCGAGATCATAAAGCGGCCACATTCACGTCGGTACTGGATGGGTTCGCGTGGCC
ATACTGTGTTTTCATTGCGTGGGTGCTTCTGCTGTTACTGCGACGCAGATTCTGTAGGCAAGGCGCAGGCTCTCT
TCTGAGGTAGAAAACACCCCATATTAATCTGAATTC

8 / 154

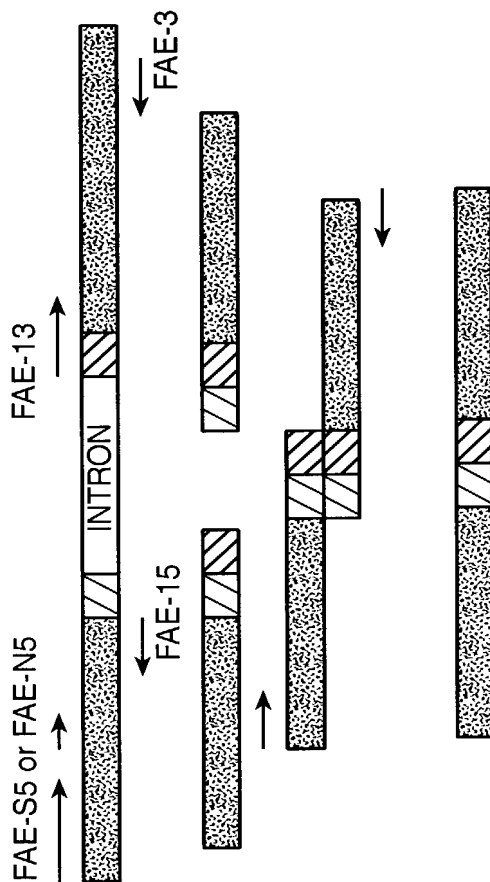


FIG. 4

FAE-13 CCGGCCACGCCCTCGGGCCCTCCCTGGGGCACTC 35-mer
 FAE-15 GGCGCCGAGGAGTGGCCGGTCAACGGTCAGCGCGTAGTCC 40-mer

intron position in original

Y A L T V T G H S L G A S L A A L
 G G A C T A C G C G T G A C C G T A C C G C C A C T C C C T C G G C G C C
 C C G G C C A C G C C C T C G G C G C C T C C C T G G C G G C A C T C
 Y A L T V T G H A L G A S L A A L
 complement, FAE-15
 FAE-13

FIG. 5

9 / 154

Vector Construction

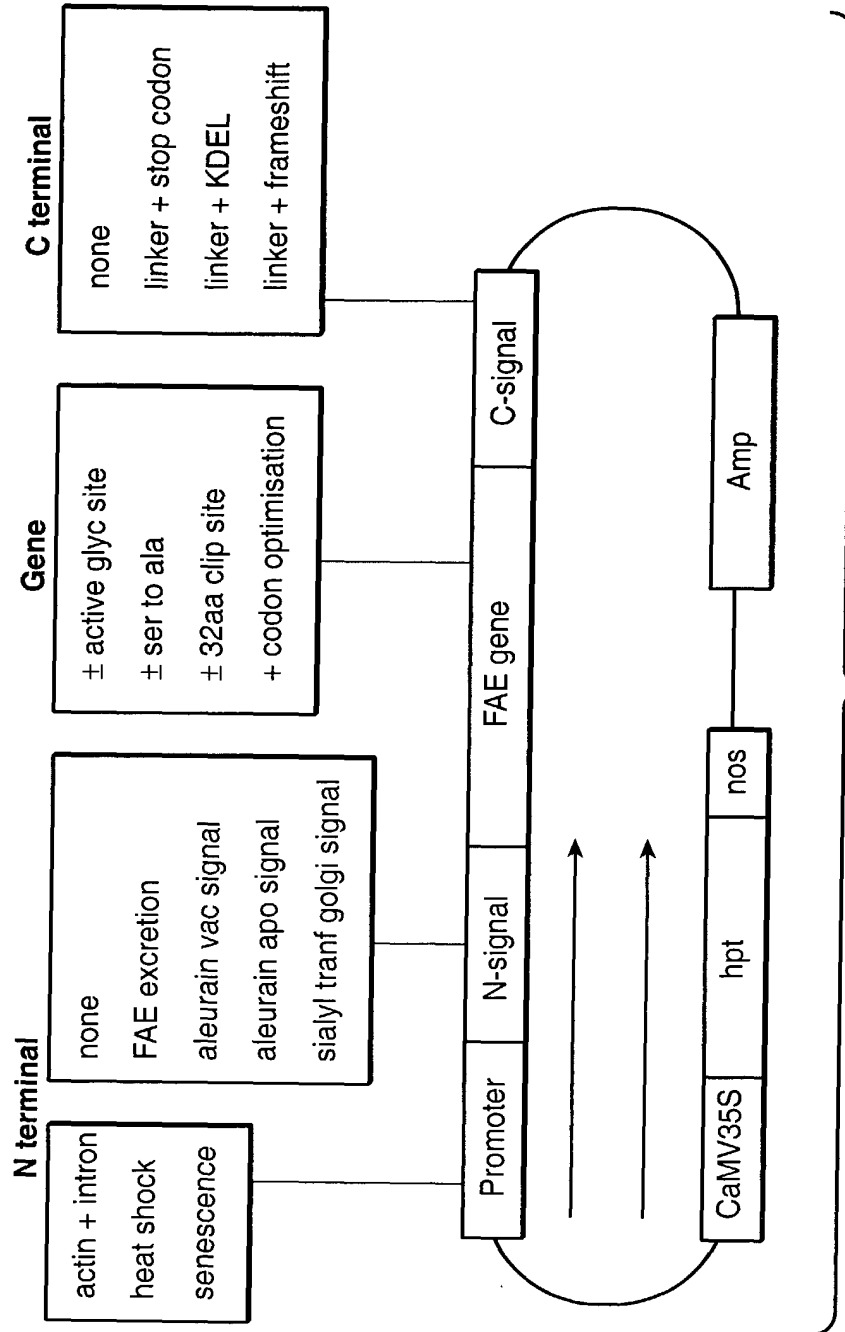


FIG._6

10 / 154

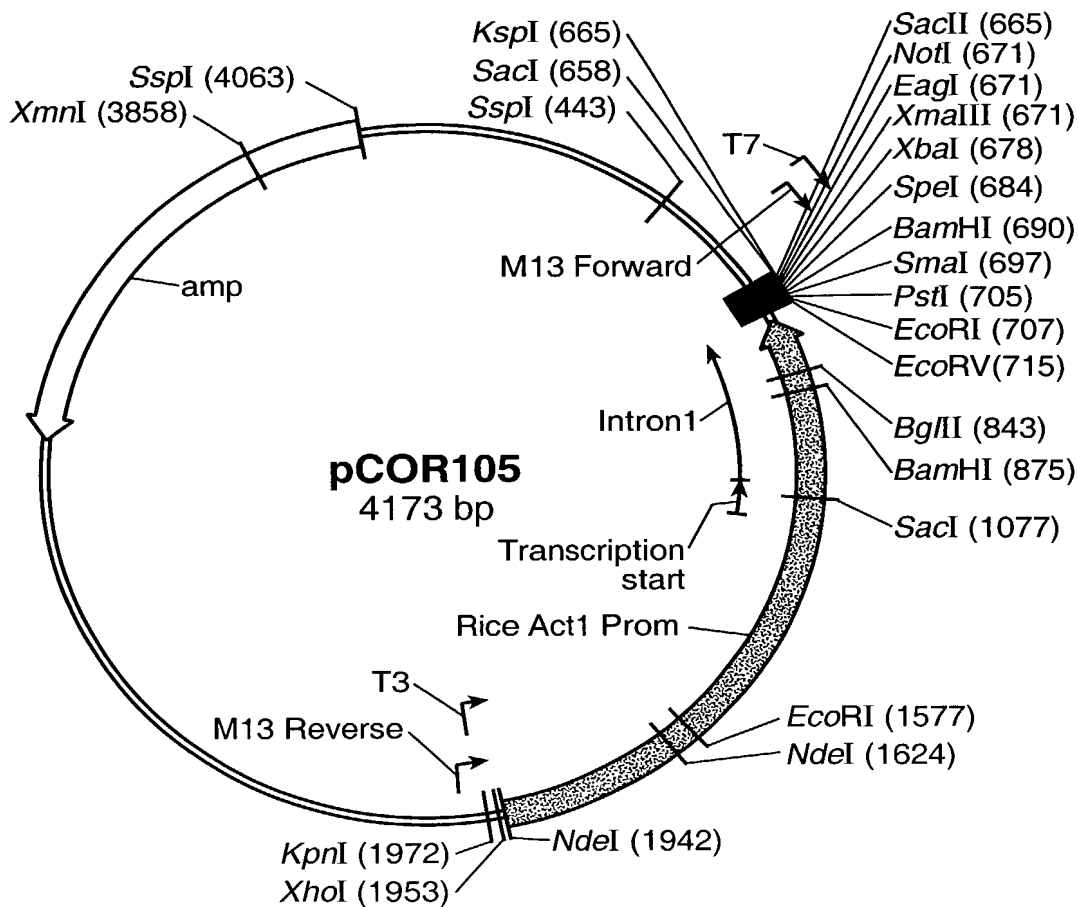


FIG. 7

11 / 154

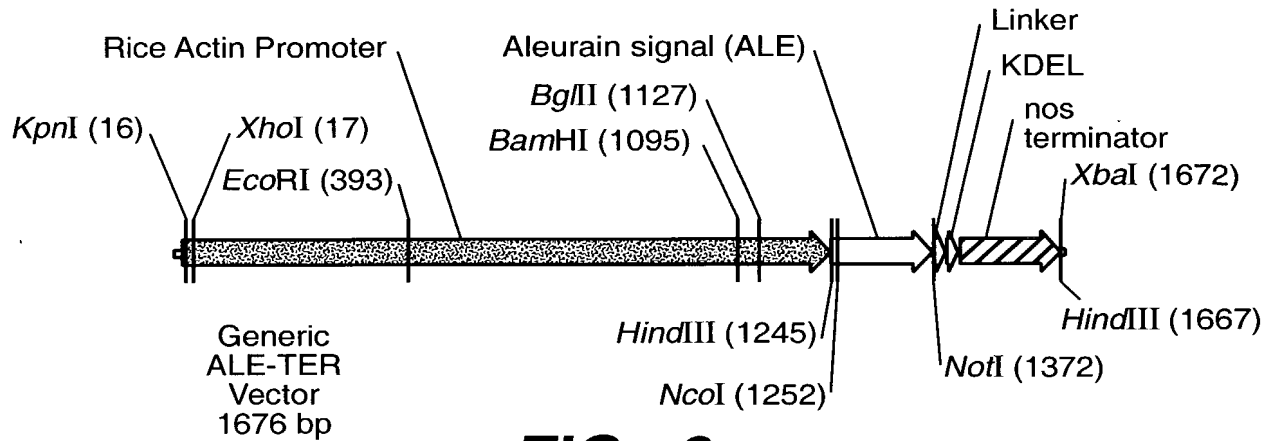


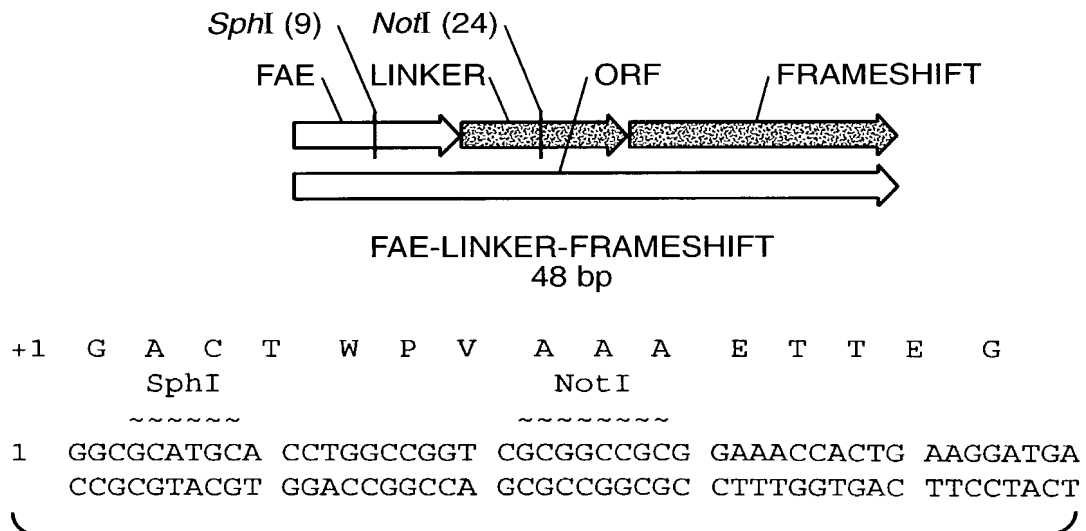
FIG._8

KDEL-COOH ER retention sequence

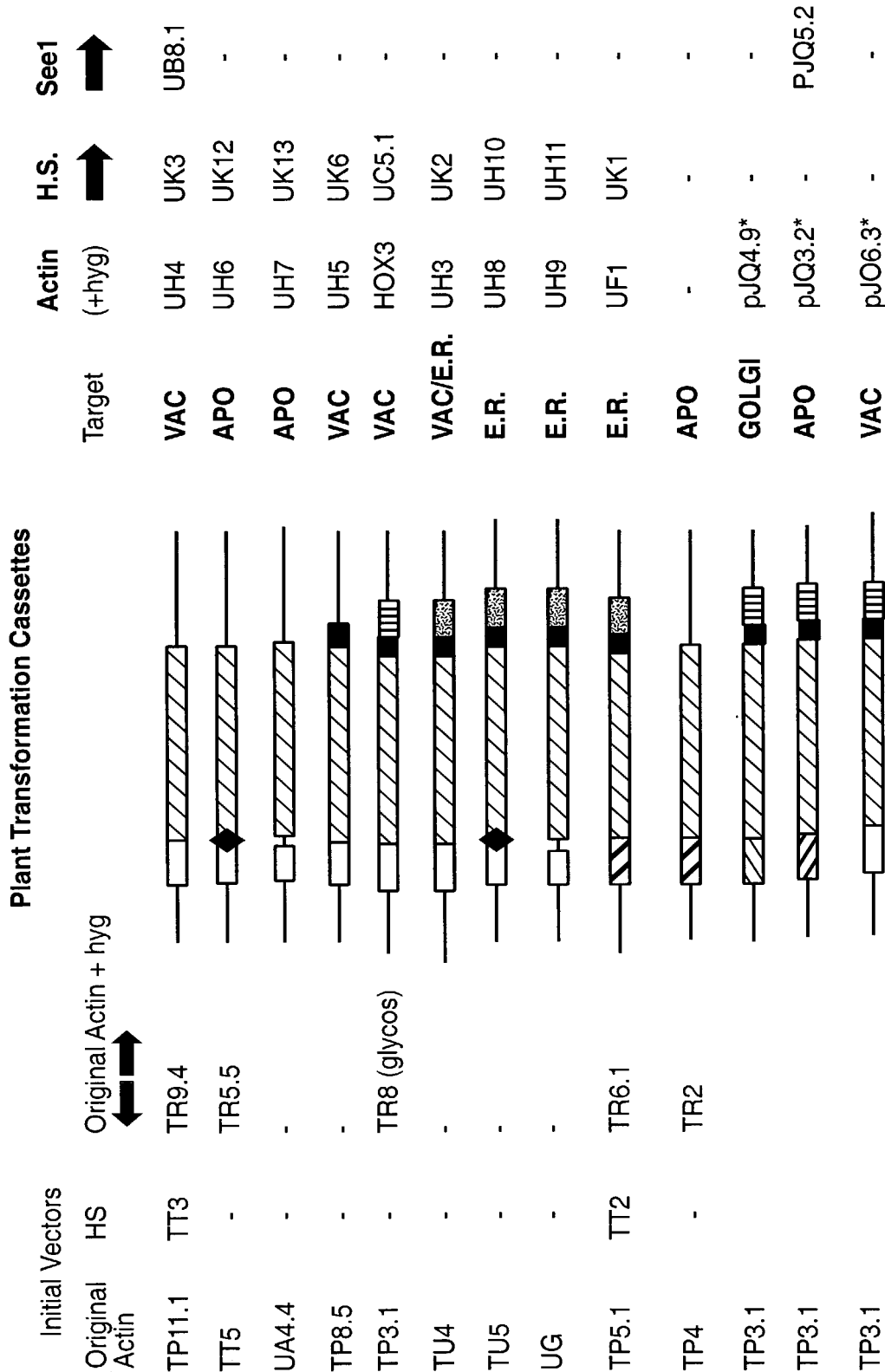
NotI
 ~~~~~  
 A A A K P L K D E L \*  
 1 GCGGCCGCGA AACCACTGAA GGATGAGCTG TAA

**FIG.\_9**

**FAE-LINKER-FRAMESHIFT Structure and Sequence**



**FIG.\_10**



\* - Modified Actin Promoter (Kpn1-EcoR1 Deletion and Restored NCO Site)

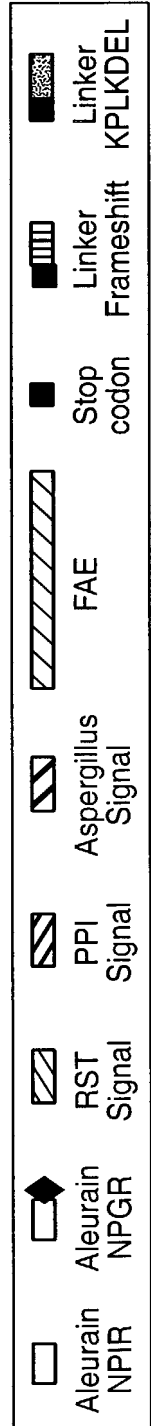


FIG.-11

## Vectors

### Original Actin promoter in pCOR105

|       | Target | Signal sequences                                                      | Vectors                                    |
|-------|--------|-----------------------------------------------------------------------|--------------------------------------------|
| (i)   | APO    | - aleurain-NPGR-FAE<br>- aleurain-delNPIR -FAE                        | pUH6, pTT5, TT5.5, pTT5.1<br>pUH7, pUA4.4, |
| (ii)  | ER     | - aleurain-NPGR-FAE-linker-KDEL<br>- aleurain-delNPIR-FAE-linker-KDEL | pTU5, pUH8,<br>pUG4, pUH9,                 |
| (iii) | VAC    | - aleurain-NPIR-FAE                                                   | pTP11.1, pTR9.4, pUH4, pUK3,               |
| (iv)  | ER/VAC | - aleurain-NPIR-FAE-linker-KDEL                                       | pTU4, pUH3,                                |
| (v)   | VAC    | - aleurain-NPIR-FAE-linker-frameshift                                 | pUA1K3, pTP3.1, pUC5.11                    |
| (vi)  | VAC    | - aleurain-NPIR-FAE-linker-stop                                       | pTP8.5, pUH5                               |
| (vii) | ER     | - Aspergillus signal -FAE-KDEL                                        | pTP5.1, pTP6.1, pUF1,                      |

### Modified actin promoter (Kpn1-EcoR1 deletion and restored NCO site)

|       |       |                                       |        |
|-------|-------|---------------------------------------|--------|
| (i)   | VAC   | - aleurain-NPIR-FAE-linker-frameshift | pJ06.3 |
| (ii)  | GOLGI | - RST-FAE-linker-frameshift           | pJQ3.2 |
| (iii) | APO   | - PPI-FAE-linker-frameshift           | pJQ4.9 |

### Heat-shock promoter

|       |        |                                                                           |                                    |
|-------|--------|---------------------------------------------------------------------------|------------------------------------|
| (i)   | APO    | - aleurain-NPGR-FAE<br>- aleurain-delNPIR-FAE<br>- Aspergillus signal-FAE | pUH12<br>pUH13<br>pTP4a2, pTR2.22, |
| (ii)  | ER     | - aleurain-NPGR-FAE-linker-KDEL<br>- aleurain-delNPIR-FAE-linker-KDEL     | pUH10<br>pUH11                     |
| (iii) | VAC    | - aleurain-NPIR -FAE                                                      | pUK3, pTT3                         |
| (iv)  | ER/VAC | - aleurain-NPIR-FAE-linker-KDEL                                           | pUK2                               |
| (v)   | VAC    | - aleurain-NPIR-FAE-linker-frameshift                                     | pUC5.11, pHOX3                     |
| (vi)  | VAC    | - aleurain-NPIR-FAE-linker-stop                                           | pUK6                               |
| (vii) | ER     | - Aspergillus signal -FAE-KDEL                                            | pUK1, pTT2                         |

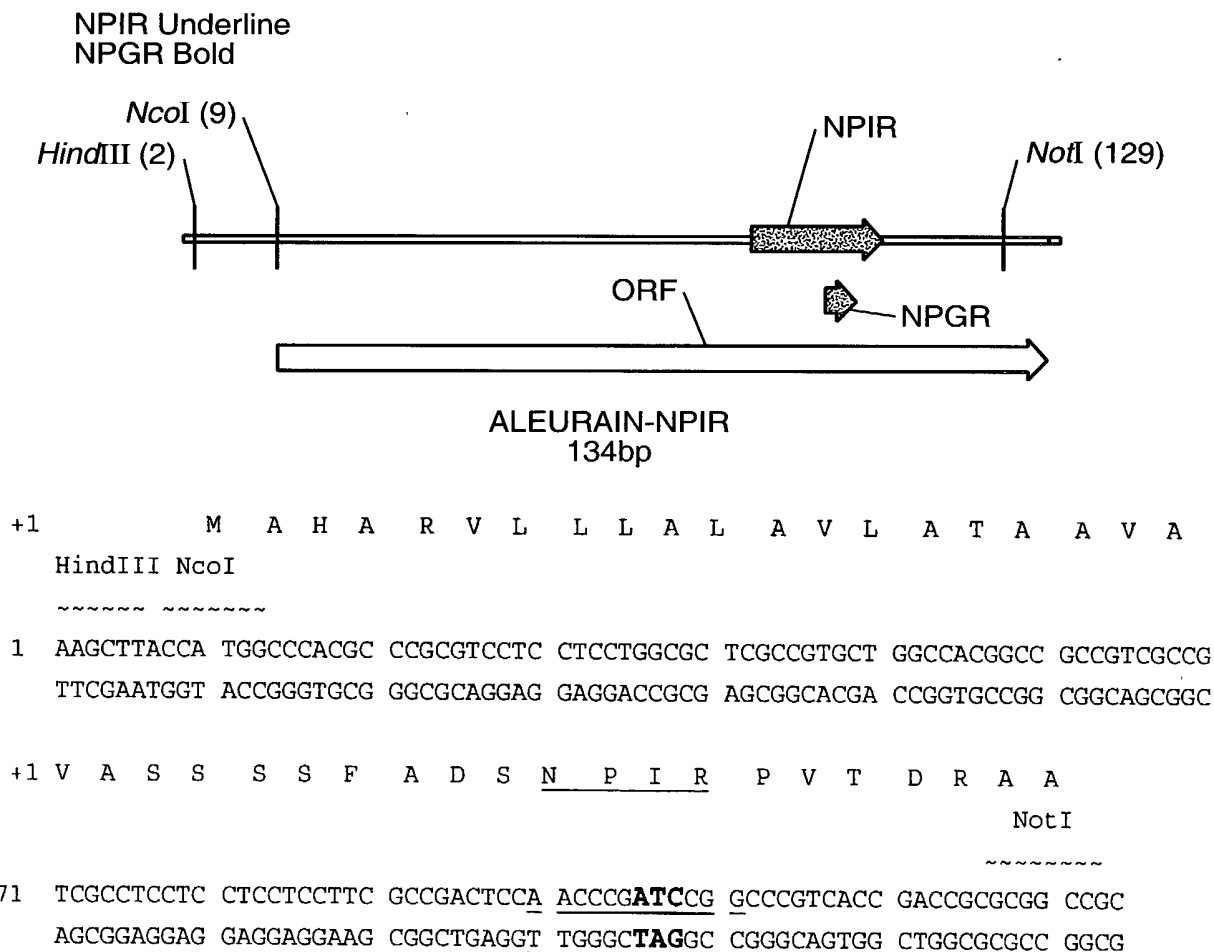
### Senescence promoter

|      |     |                                  |        |
|------|-----|----------------------------------|--------|
| (i)  | APO | - Sec1-PPI-FAE-linker-frameshift | pJQ5.2 |
| (ii) | VAC | - Sec1-aleurain-deleted NPIR-FAE | pUB8.1 |

## FIG. 12

14 / 154

# **ALEURAIN-NPIR (Vacuolar) and NPGR (Apoplast) Structure and Sequence**



**FIG. 13**

15 / 154

RAT SIALYL TRANSFERASE Golgi signal sequence

HindIII  
~~~~~  
M I H T N L K K K F S L F I L V F L L F A
1 AAGCTTACCA TGATCCACAC CAACCTCAA AAGAAGTTCT CCTCTTTCAT CCTCGTCTTC CTCTCTTCG
· V I C V W K K G S D Y E A L T L Q A K E F Q M
71 CCGTGATCTG CGTGTGGAAG AAGGGCTCCG ACTACGAGGC CCTCACCTC CAAGCCAAGG AGTTCCAAAT
NotI
~~~~~  
· A A  
141 GGCGGCCCGC

FIG.\_14

POTATO PROTEASE INHIBITOR II Apoplast signal sequence

HindIII  
~~~~~  
M X V H K E V N F V A Y L L I V L G L L L
1 AAGCTTACMA TGGMCGTGCA CAAGGAGGTS AACTTCGTSG CCTACCTCCT GATCGTCTC
GGCCTCCTCT
NcoI
~~~~~  
· L V S A M E H V D A K A C T X E C G N L  
G F G .  
71 TGCTCGTSTC CGCCATGGAG CACGTGGAGC CCAAGGCCTG CACCCKGAG TCGGCAACC  
TCGGCTTCGG  
NotI  
~~~~~  
· I C P A A A
141 CATCTGCCCG GCGGCCGC

FIG._15

16 / 154

Targeting Expression of gfp to Different Cell Compartments

Actin Promoter Targeting Vectors

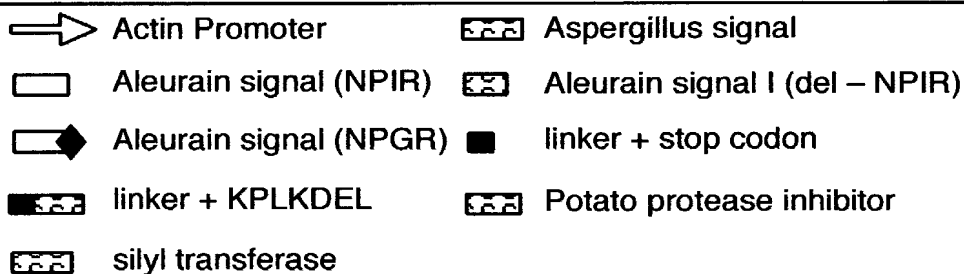
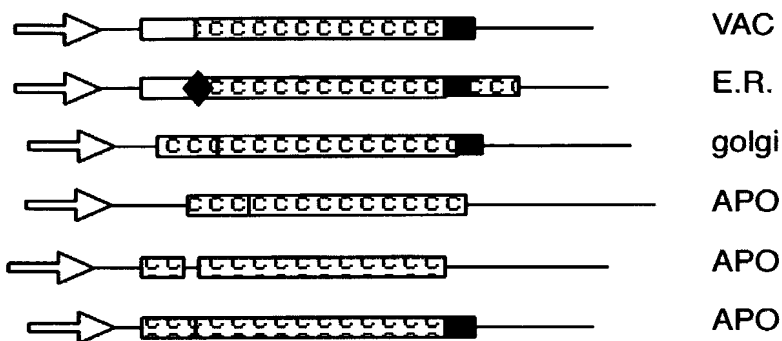


FIG._16A

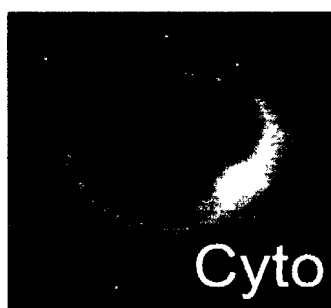


FIG._16B

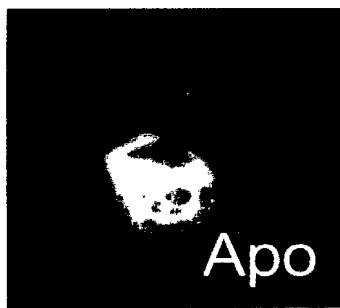


FIG._16C

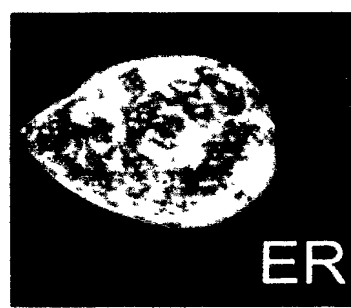


FIG._16D



FIG._16E

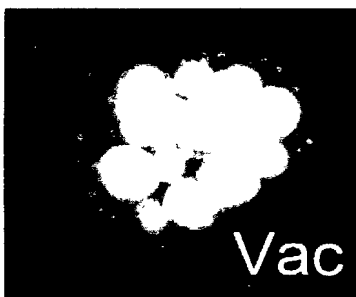


FIG._16F

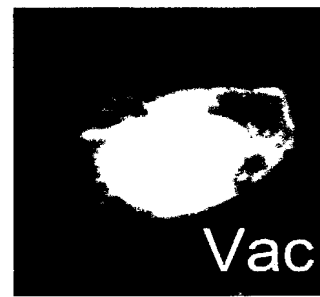


FIG._16G

FAE Activity in Transgenic *Festuca arundinacea* Leaves of Different Ages Under ER and APO Targeting Sequence

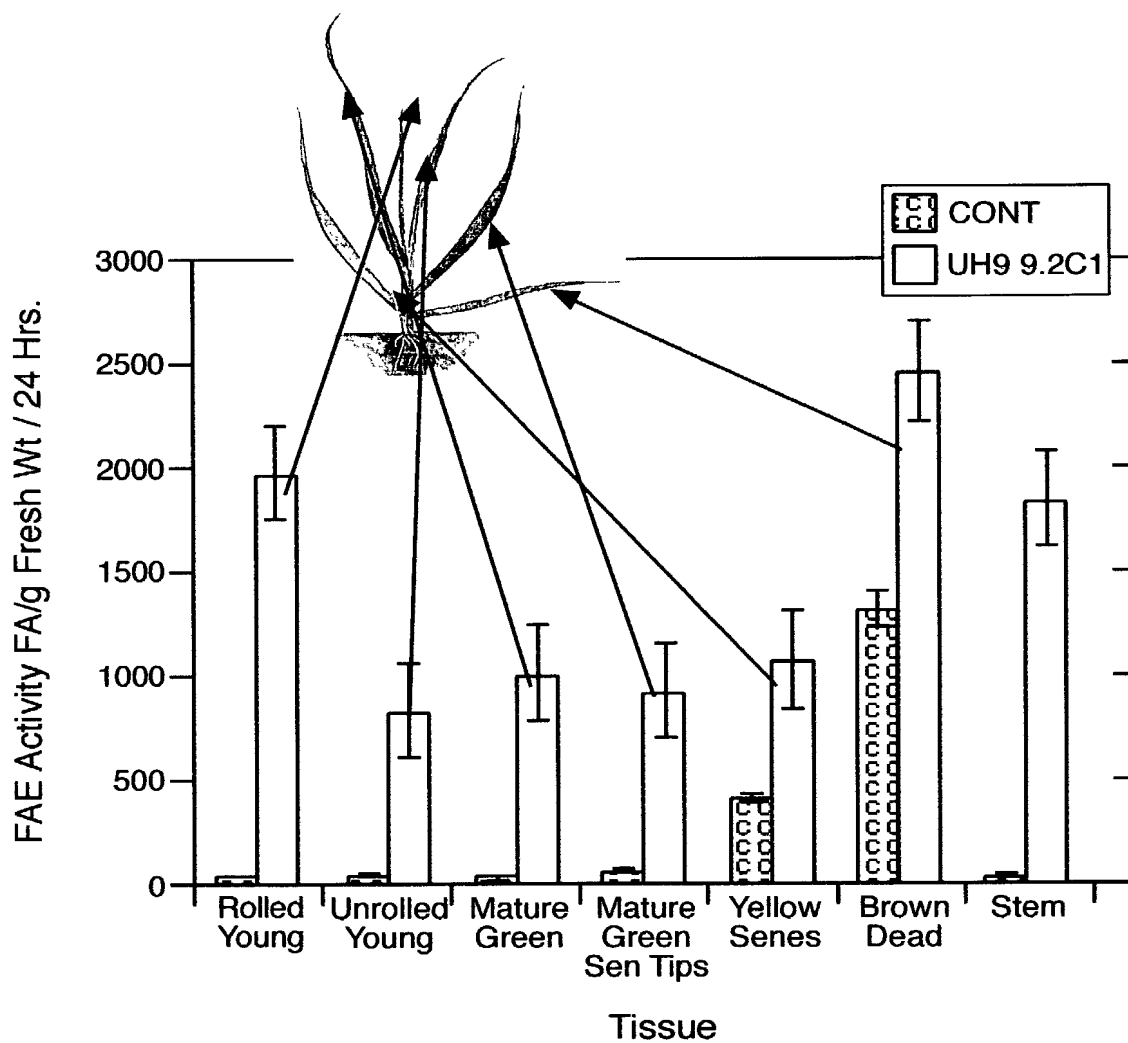


FIG. 17A

18 / 154

FAE Activity in Transgenic *Festuca arundinacea* Leaves of Different Ages Under ER and APO Targeting Sequence

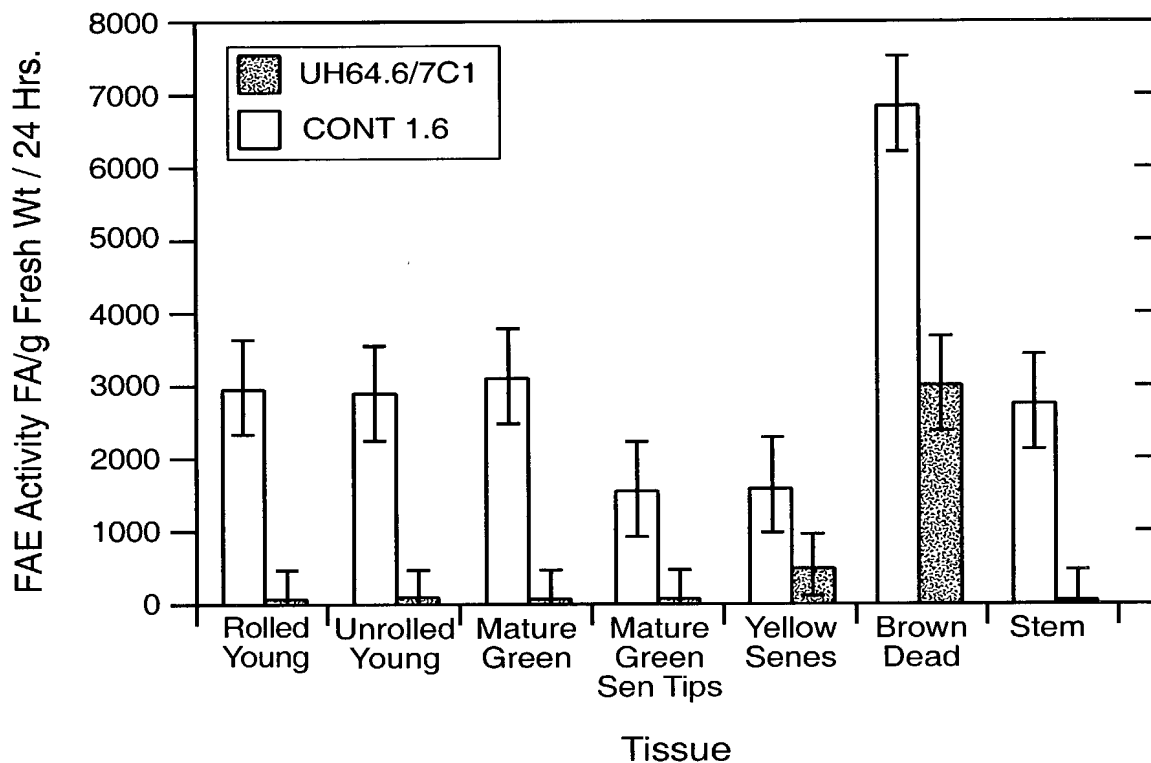
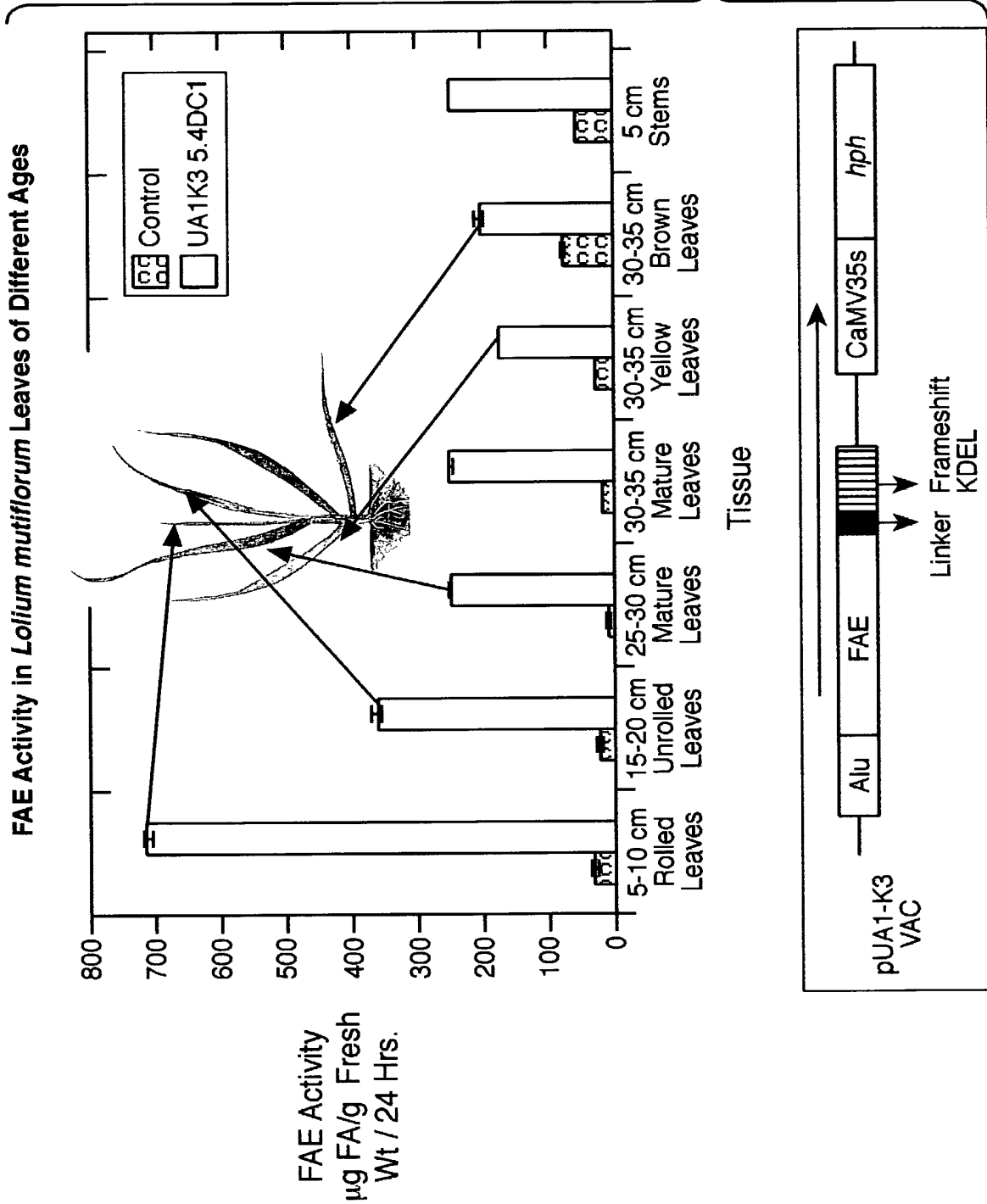


FIG._17B



FIG. 19



21 / 154

FAE Activity in Leaves of Primary Transformants of *Lolium multiflorum* Under VAC APO and ER Targeting Sequence

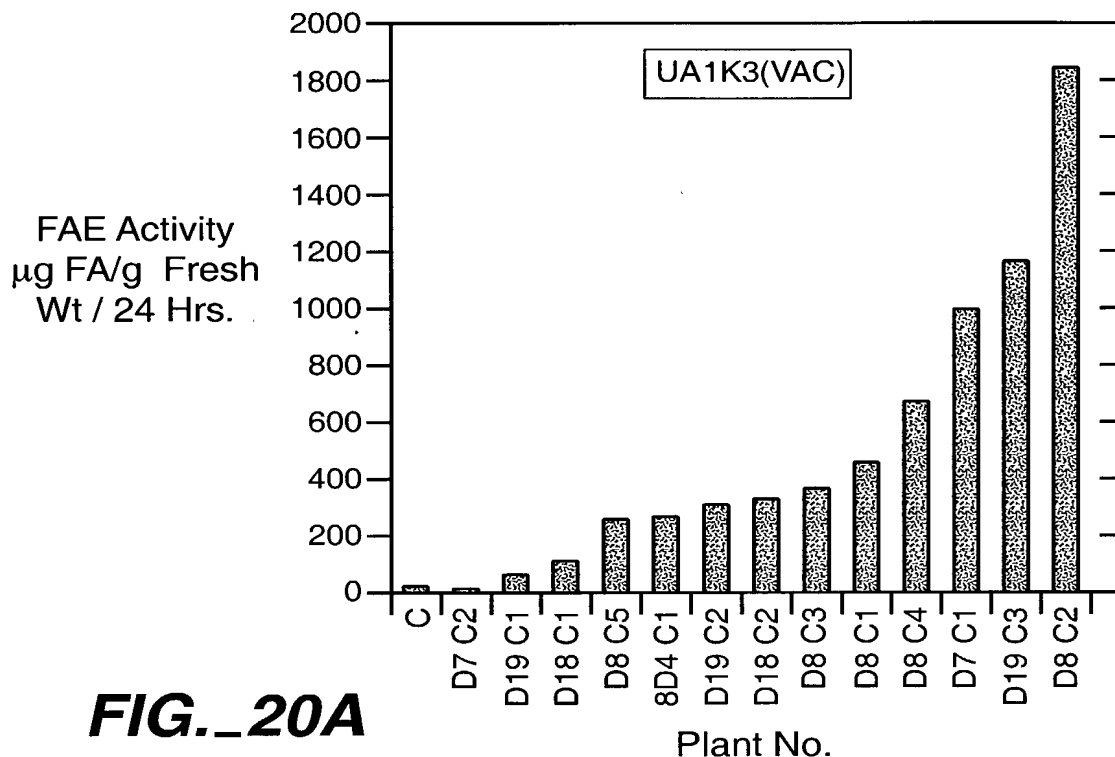


FIG._20A

FAE Activity in Leaves of Primary Transformants of *Lolium multiflorum* Under VAC APO and ER Targeting Sequence

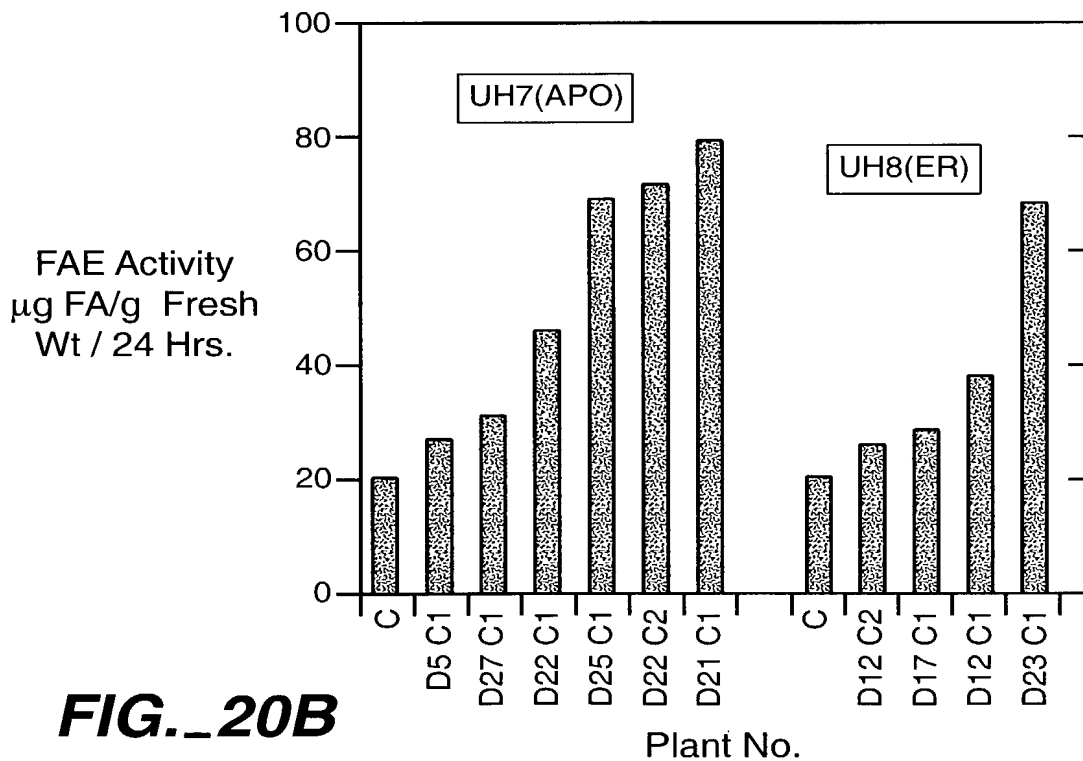


FIG._20B

22 / 154

Levels of Esterified Monomeric and Dimeric Hydroxycinnamicacids in
Festuca Arundinacea Plants Expressing FAE Under VAC Targeting Sequence

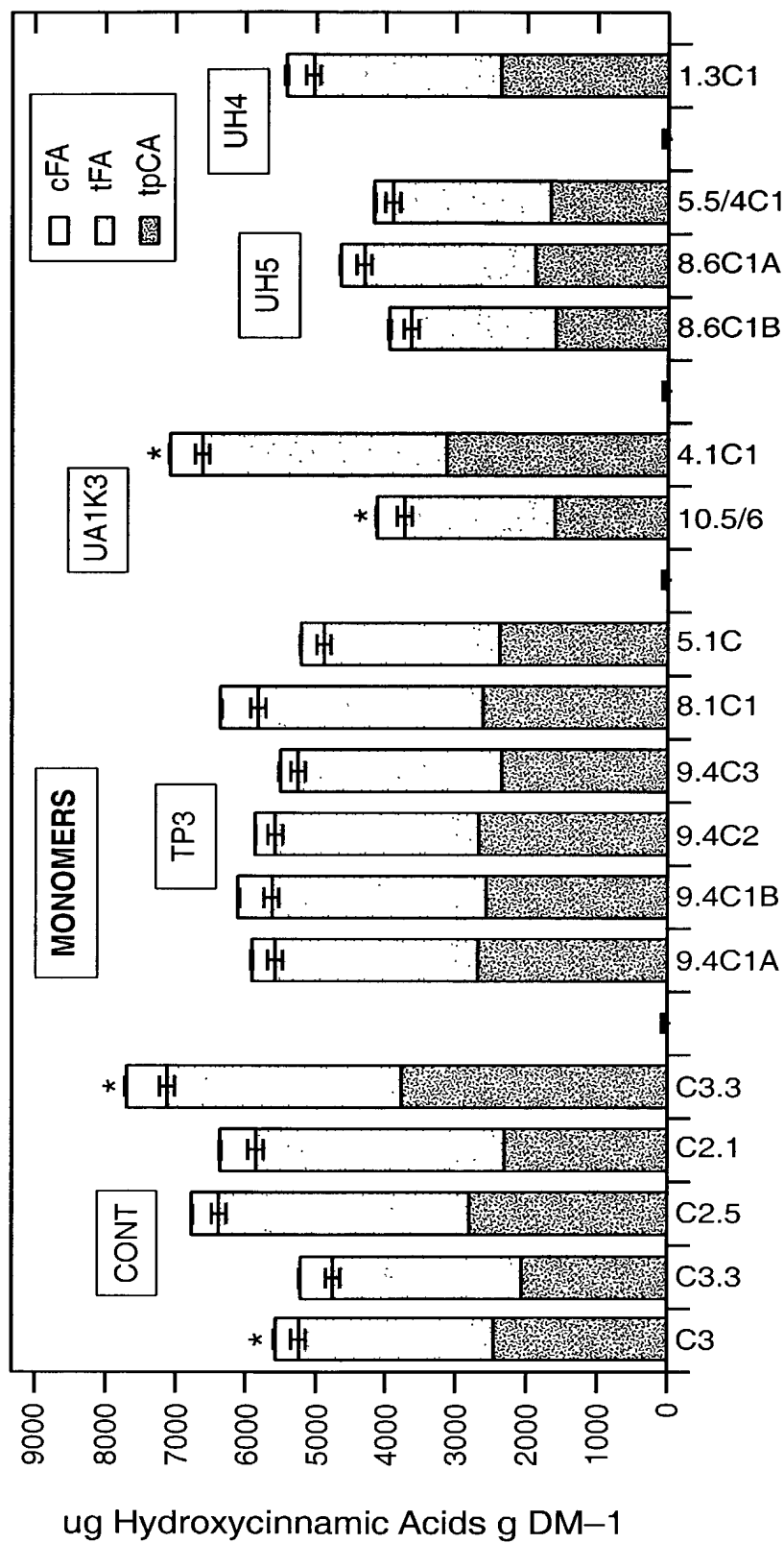


FIG..21A

23 / 154

Levels of Esterified Monomeric and Dimeric Hydroxycinnamicacids in
Festuca Arundinacea Plants Expressing FAE Under VAC Targeting Sequence

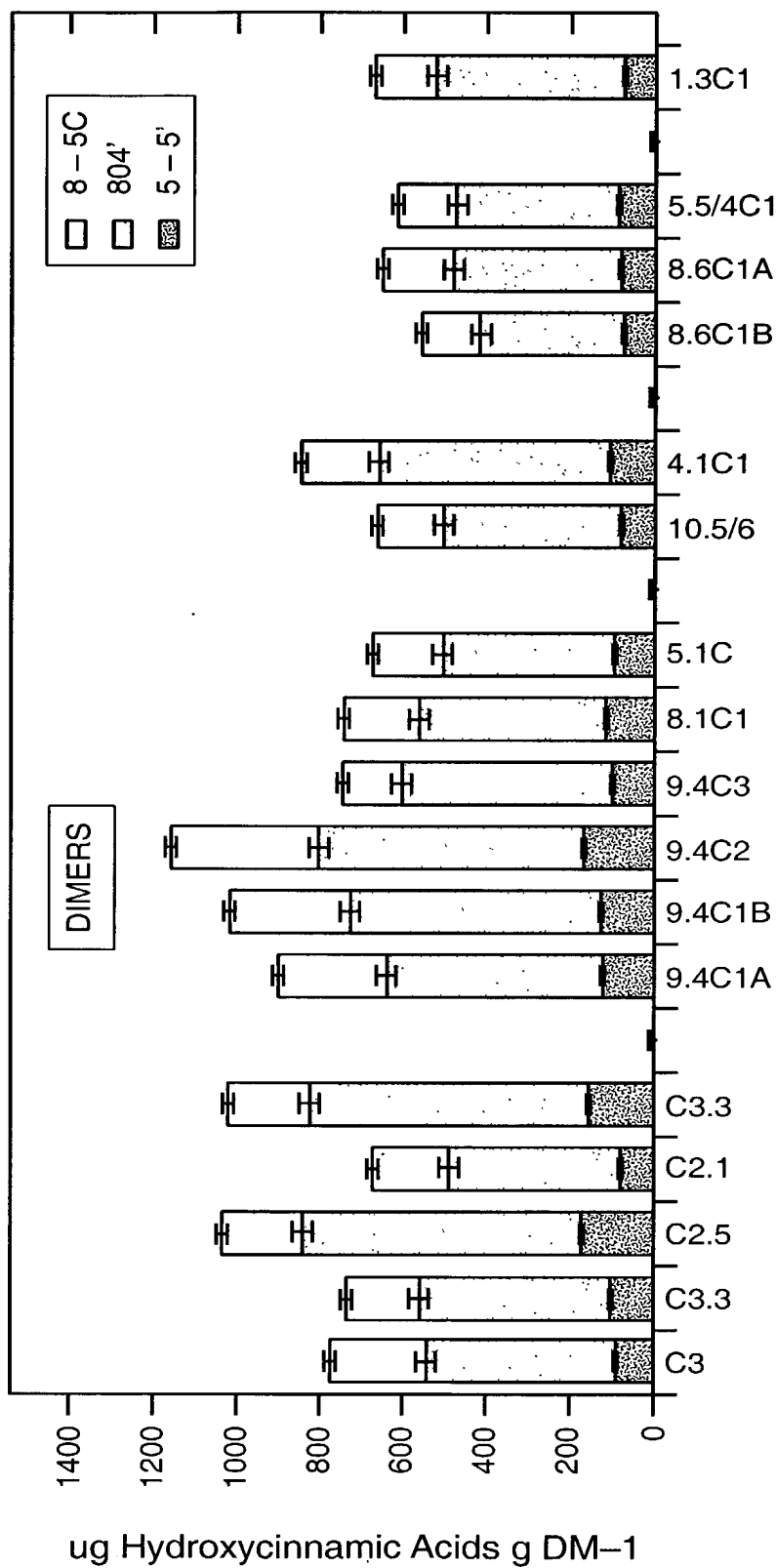


FIG..21B

24 / 154

Levels of Esterified Monomeric and Dimeric Hydroxycinnamic Acids in Leaves of *F. a.* Expressing FAE Under ER and APO Targeting Sequence

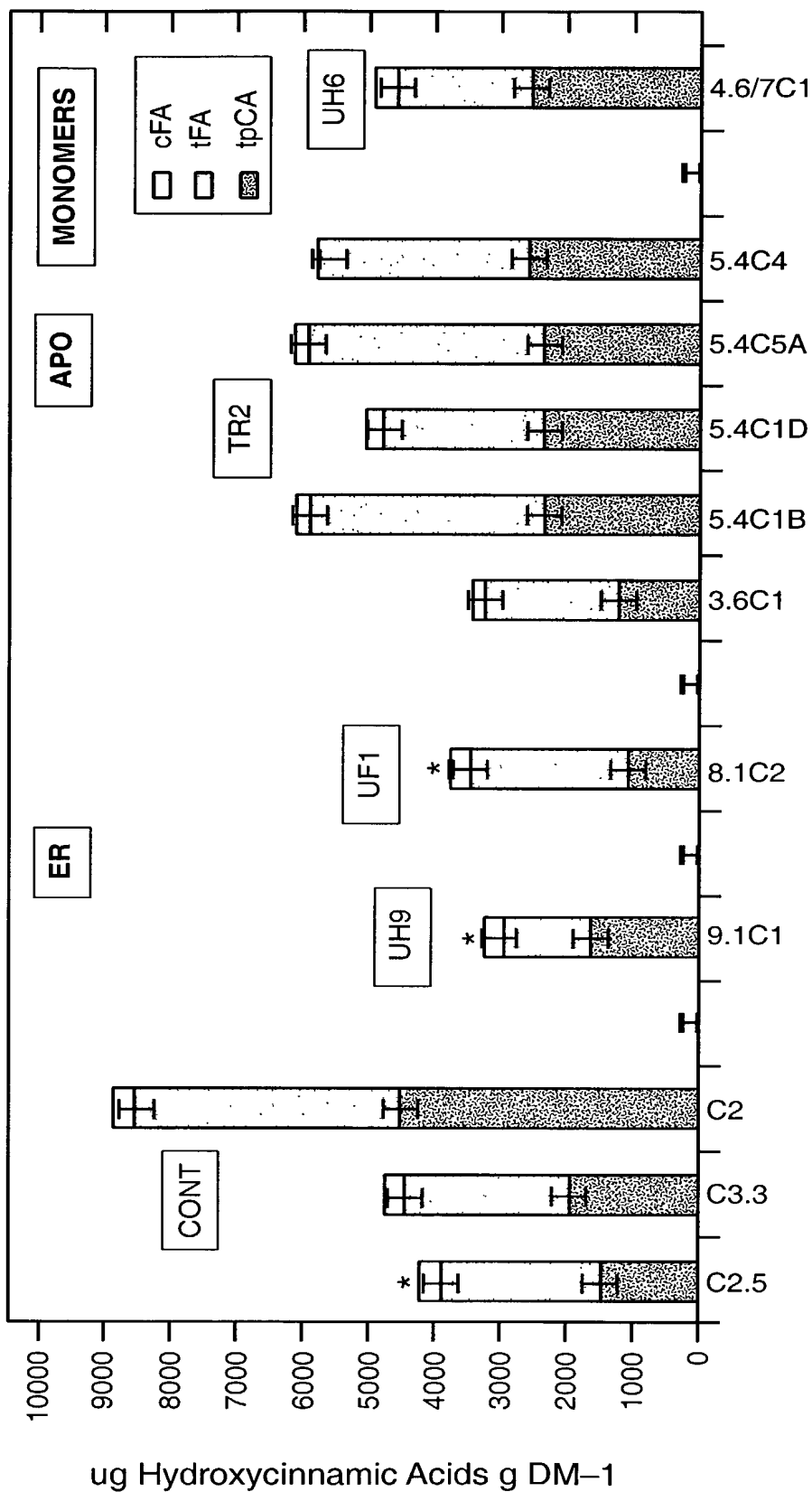


FIG. 22A

25 / 154

Levels of Esterified Monomeric and Dimeric Hydroxycinnamic Acids in Leaves of *F. a.* Expressing FAE Under ER and APO Targeting Sequence

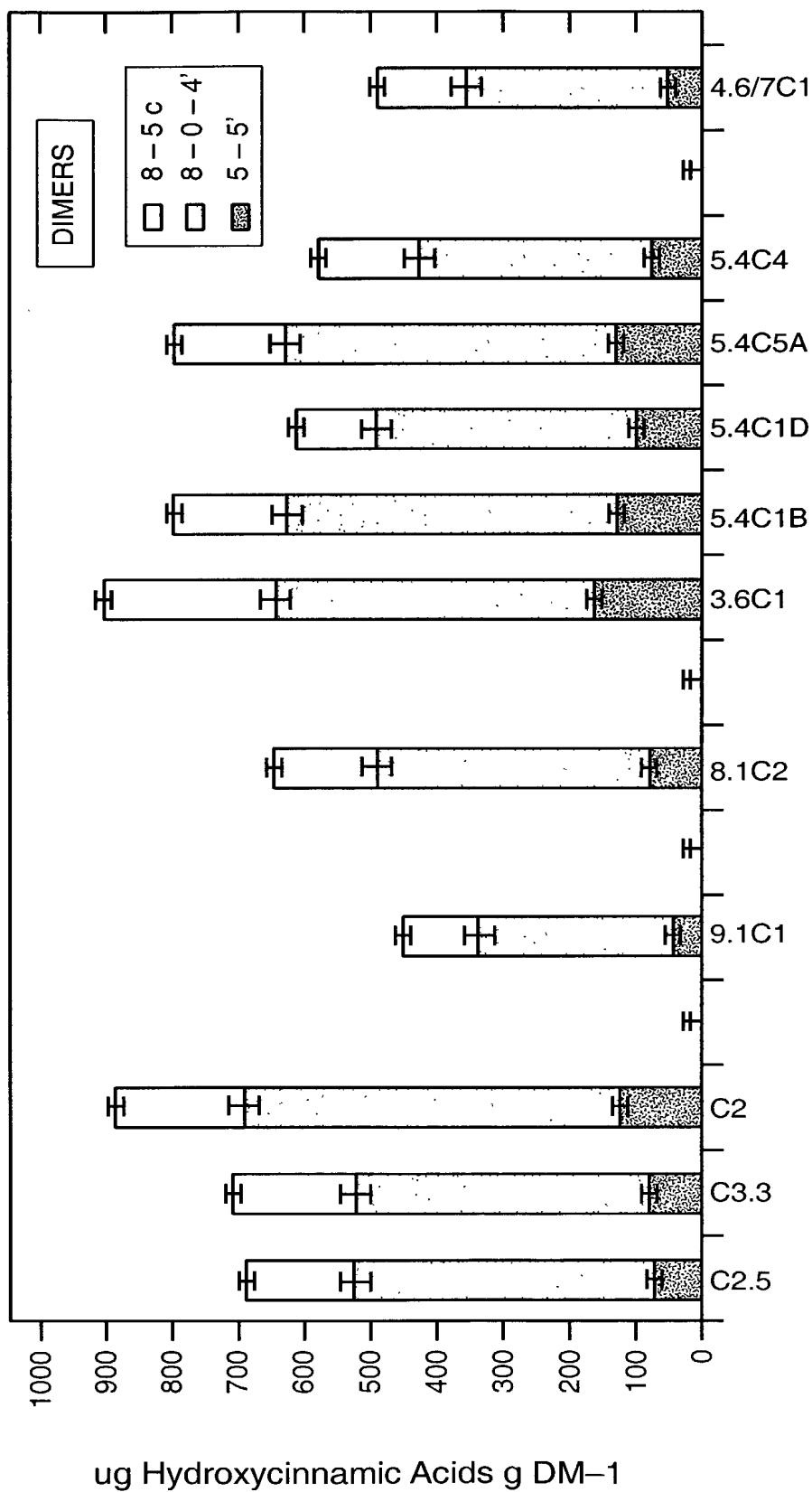


FIG.-22B

26 / 154

In Vitro Dry Matter Digestibility of Leaf Tissue of Mature
Festuca arundinacea Plants Expressing FAE Under an Actin Promoter

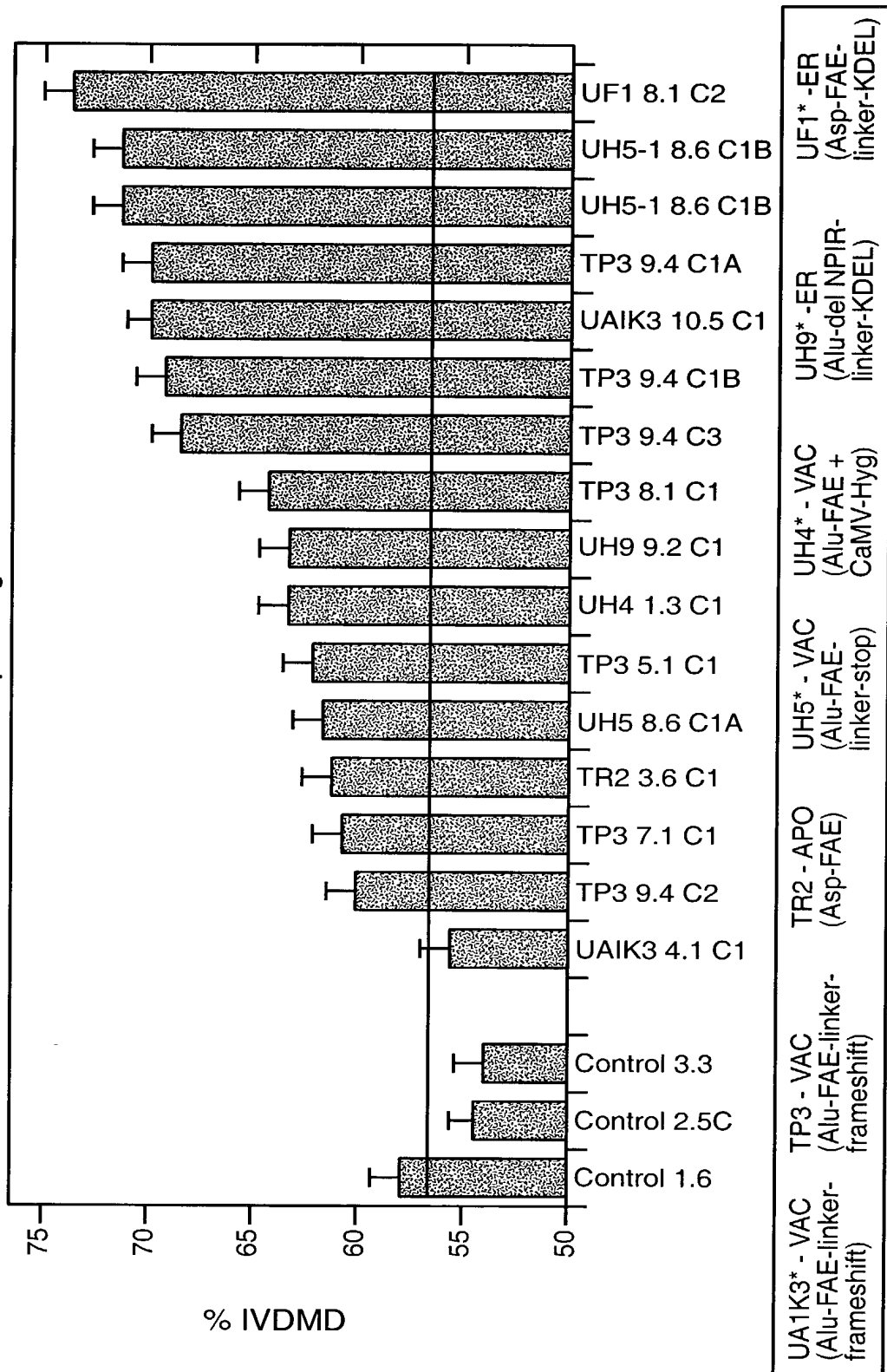


FIG._23

27 / 154

In Vitro Dry Matter Digestibility of Leaf Tissue of Mature
Lolium multiflorum Plants Expressing FAE Under an Actin Promoter

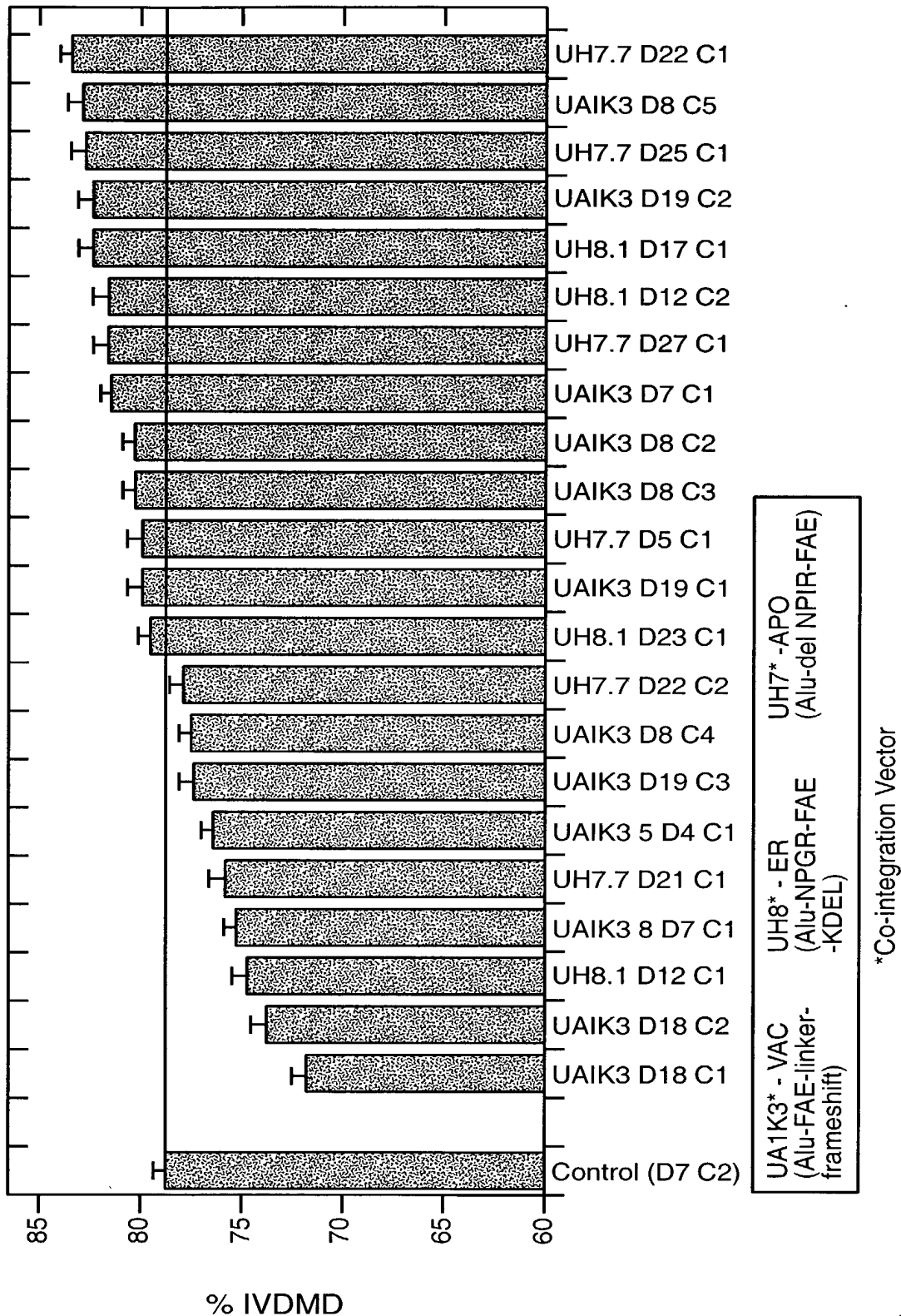


FIG._24

28 / 154

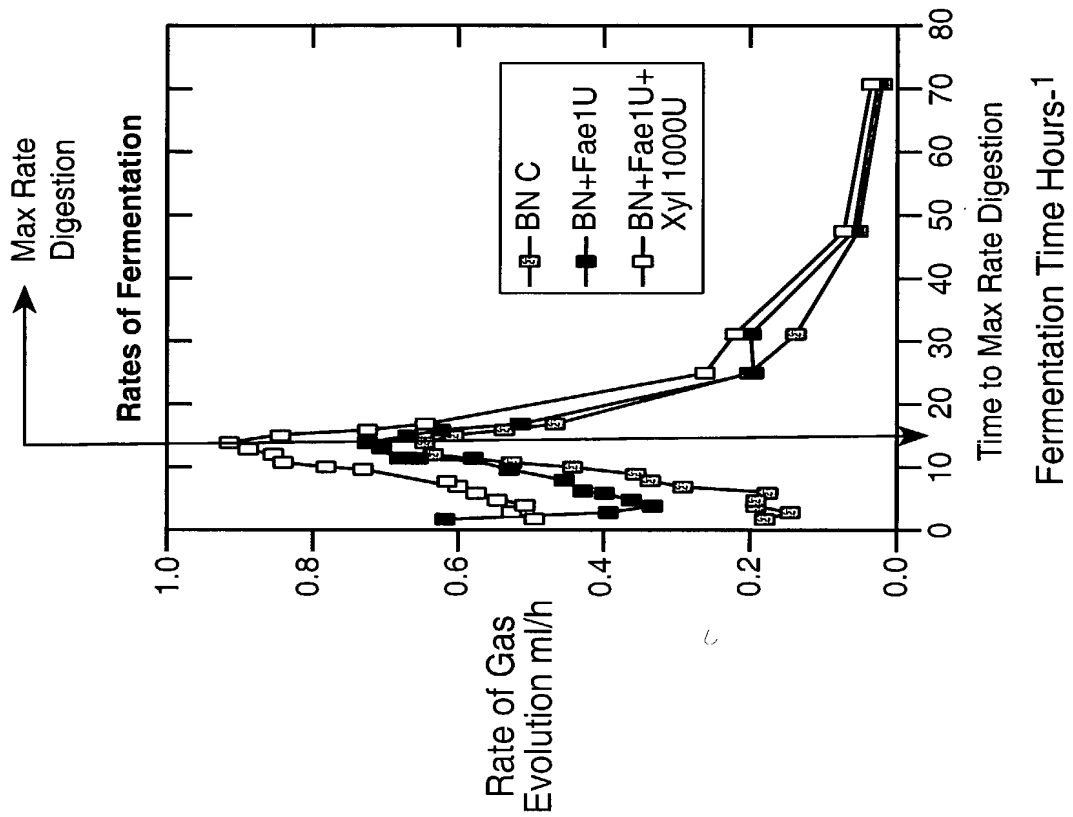


FIG._25B

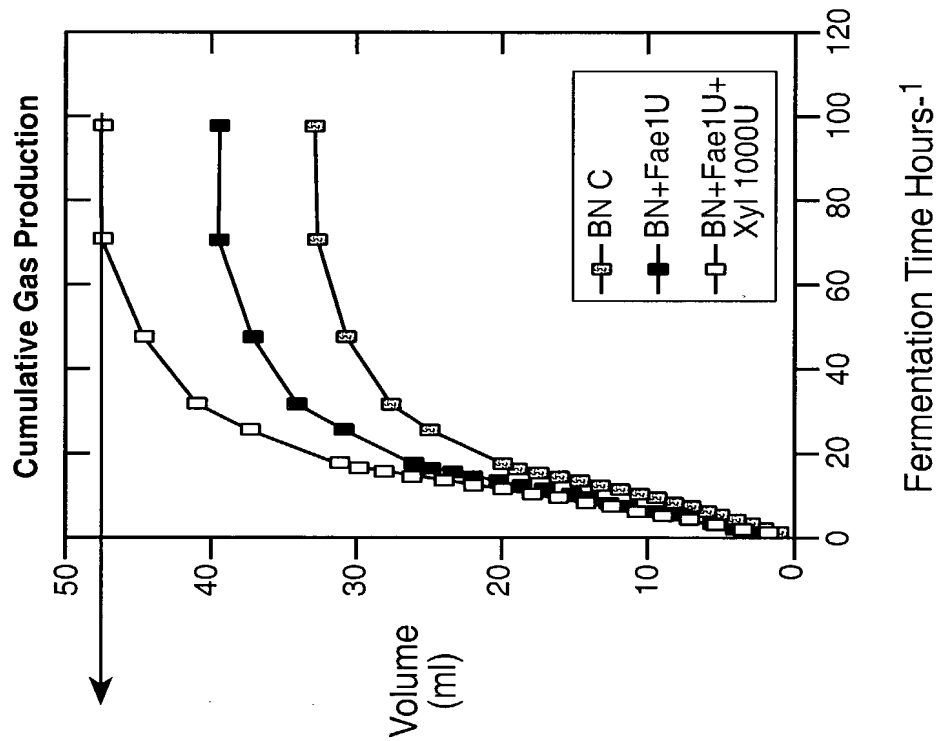


FIG._25A

29 / 154

**In-vitro Fermentation of *Festuca arundinacea* Cell Walls
 From Cell Cultures Expressing Recombinant FAE1**

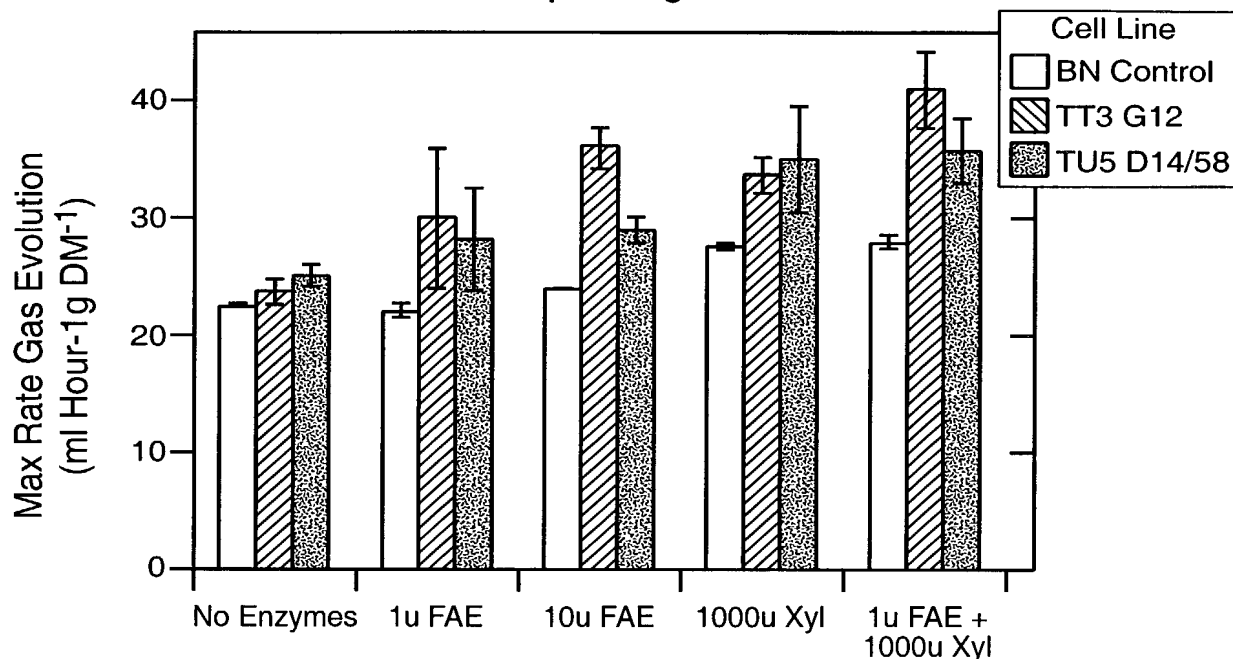


FIG._26A Maximum Rate of Digestion

**In-vitro Fermentation of *Festuca arundinacea* Cell Walls
 From Cell Cultures Expressing Recombinant FAE1**

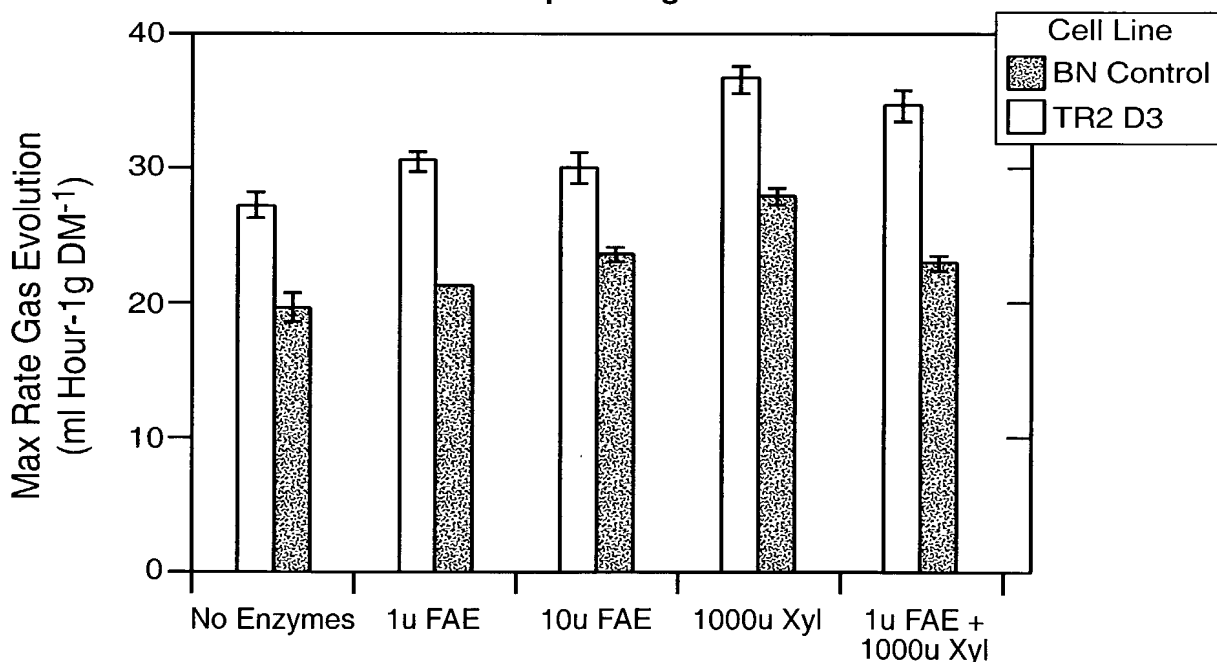


FIG._26B Maximum Rate of Digestion

30 / 154

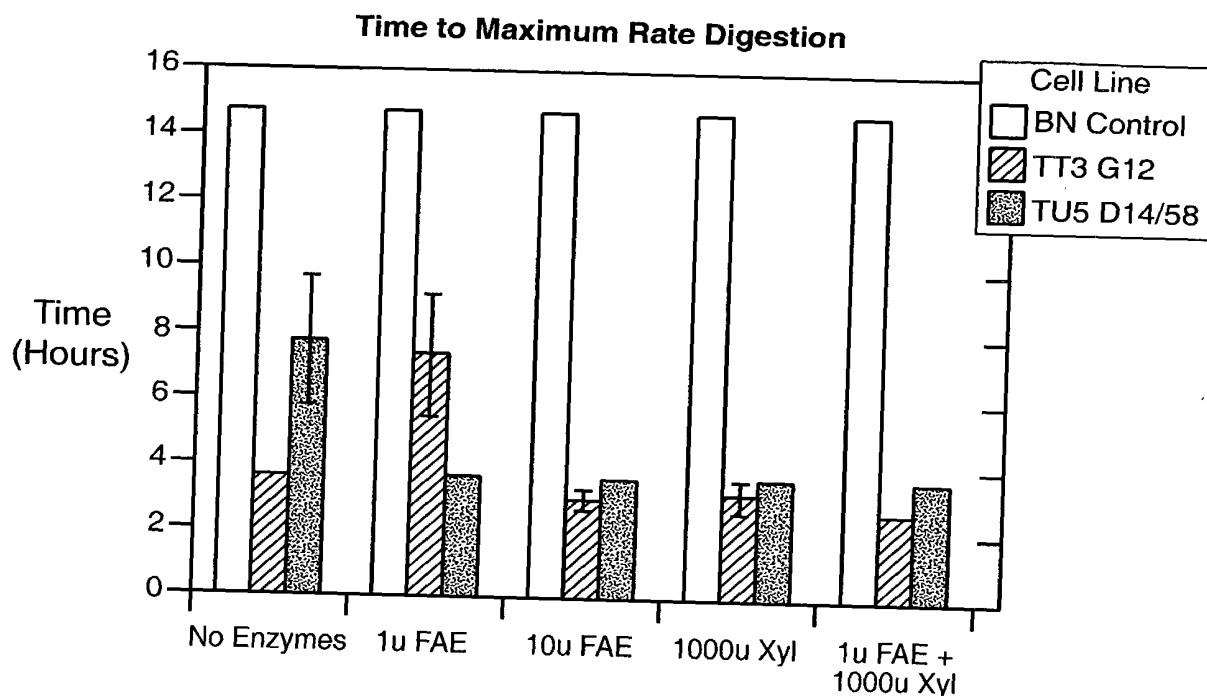


FIG._27A

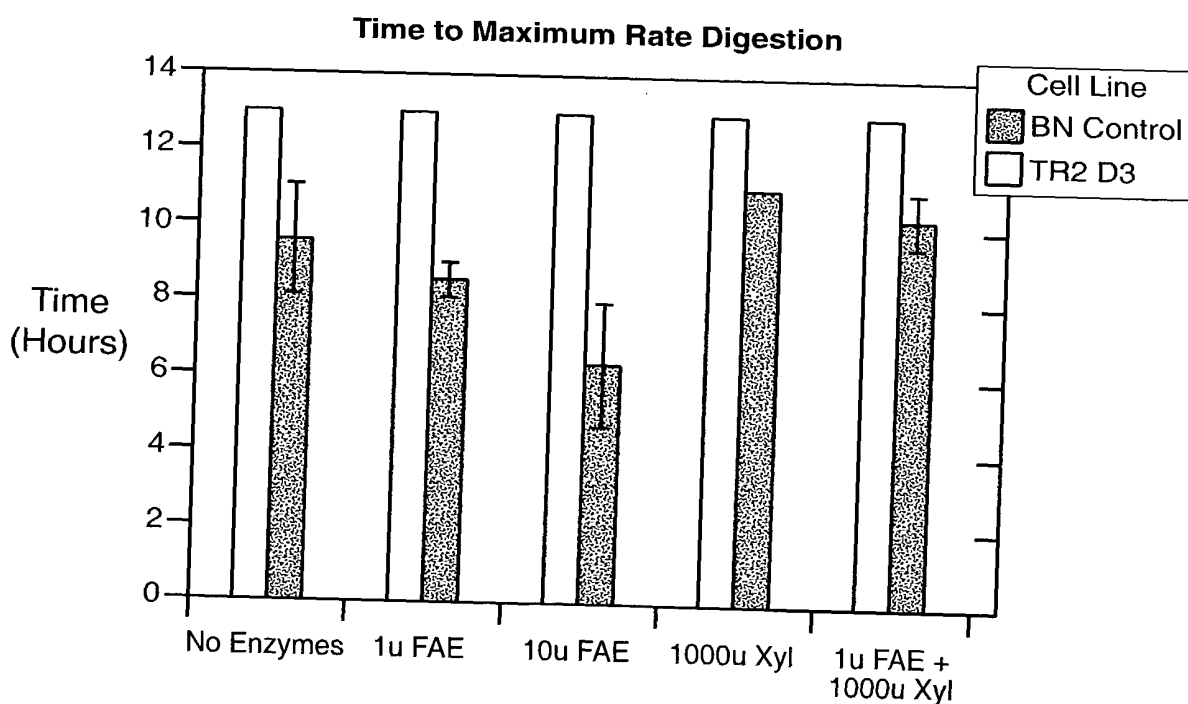


FIG._27B

31 / 154

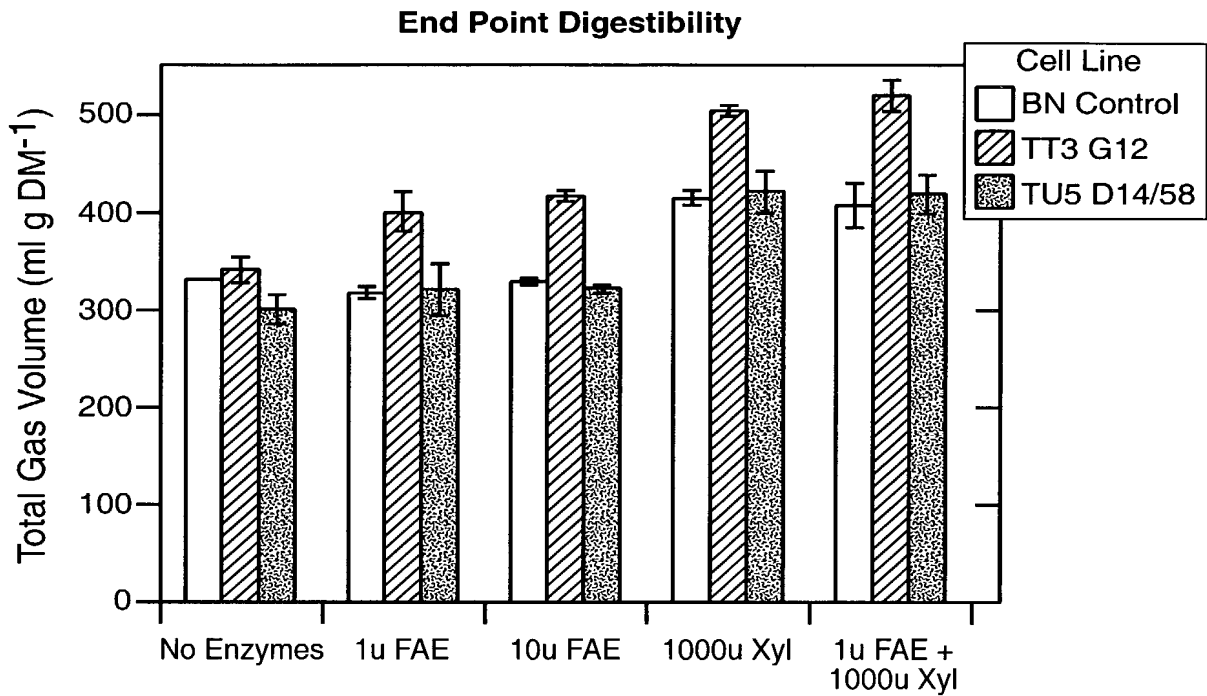


FIG._28A

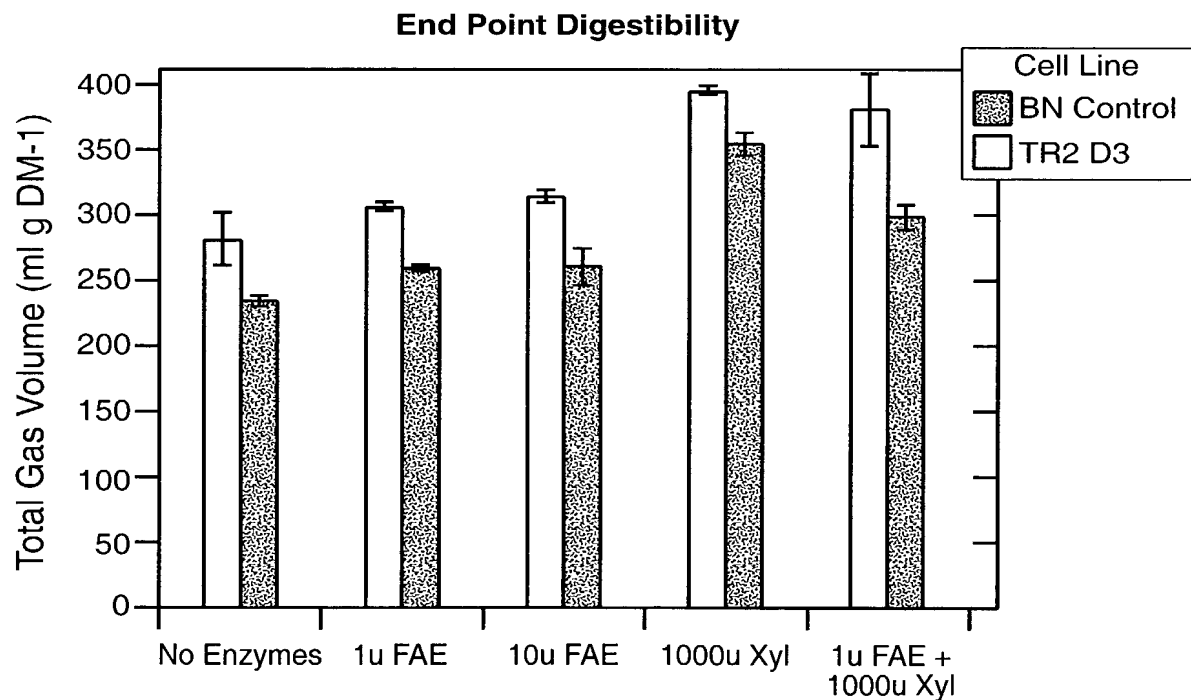


FIG._28B

32 / 154

Kinetics of FAE Activity by Ferulic Acid Release from Cell Wall under Self Digestion in *Festuca arundinacea* and Stimulation by Xylanase

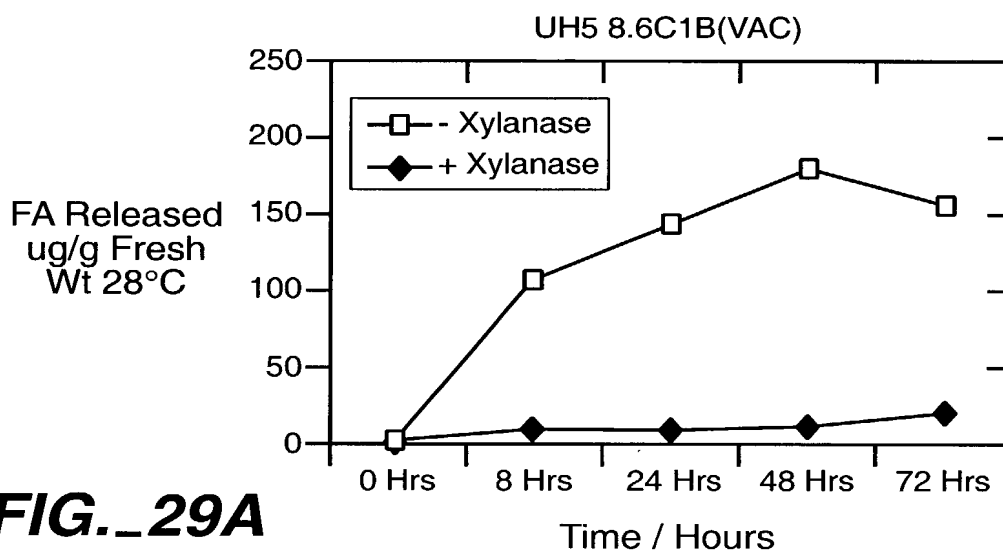


FIG._29A

Kinetics of FAE Activity by Ferulic Acid Release from Cell Wall under Self Digestion in *Festuca arundinacea* and Stimulation by Xylanase.

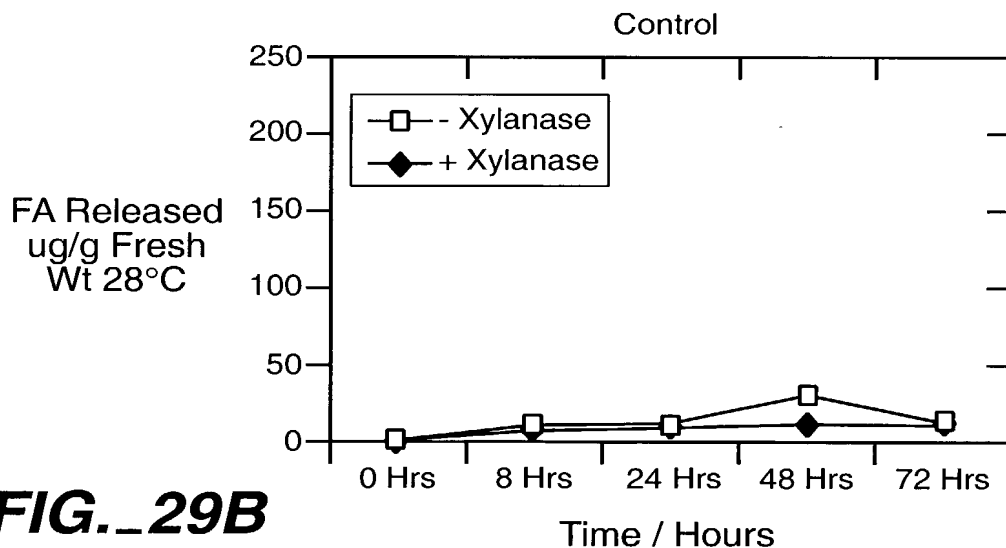


FIG._29B

33 / 154

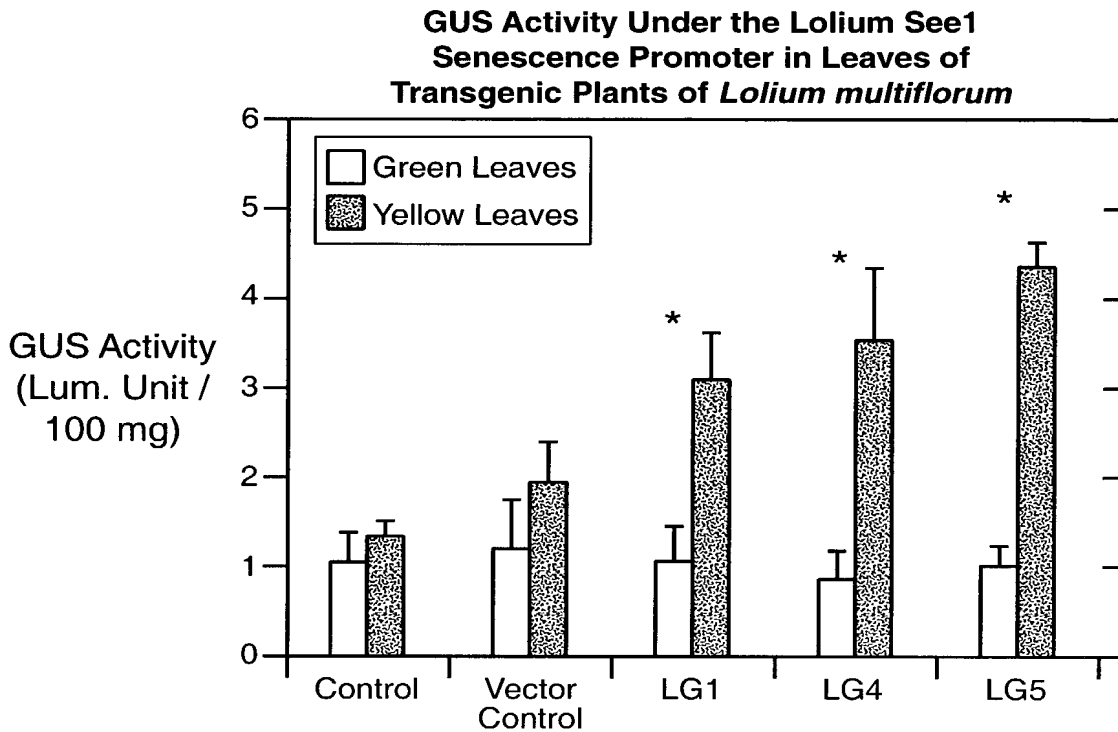


FIG._30

34 / 154

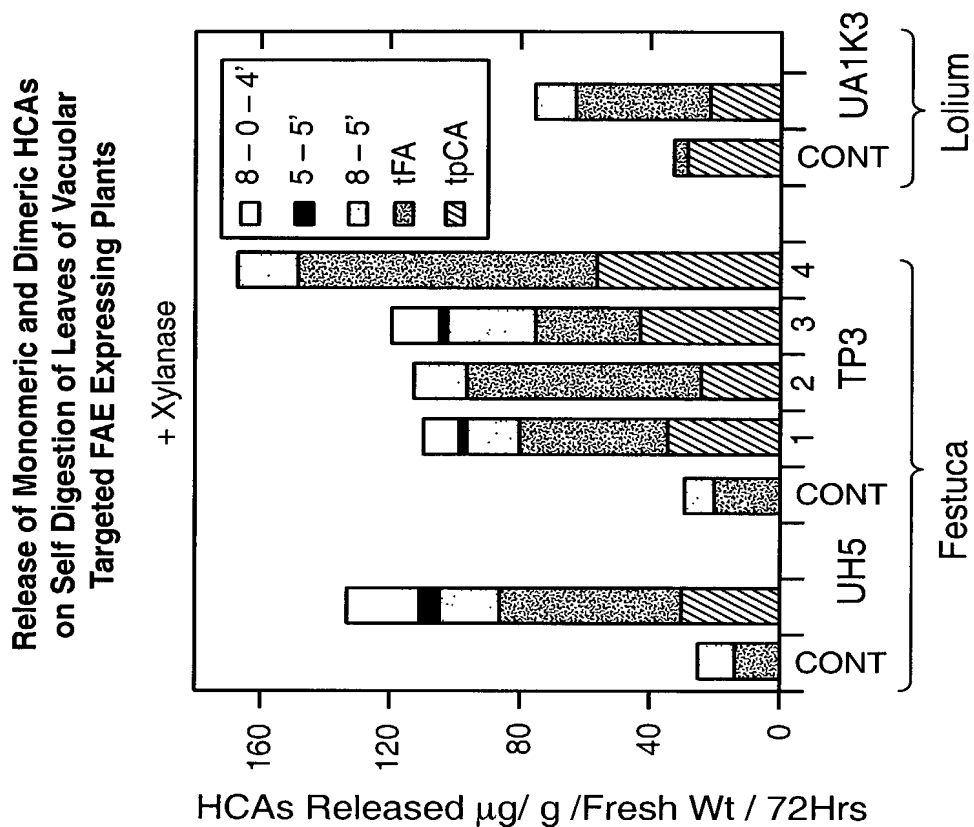


FIG._31B

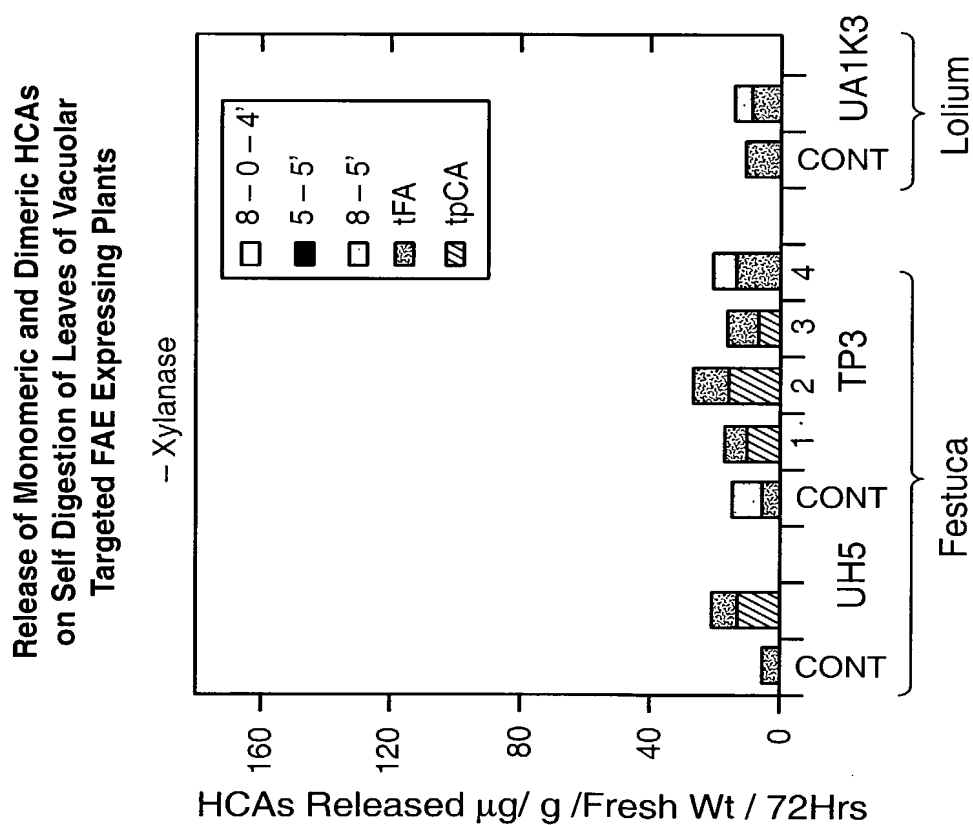


FIG._31A

35 / 154

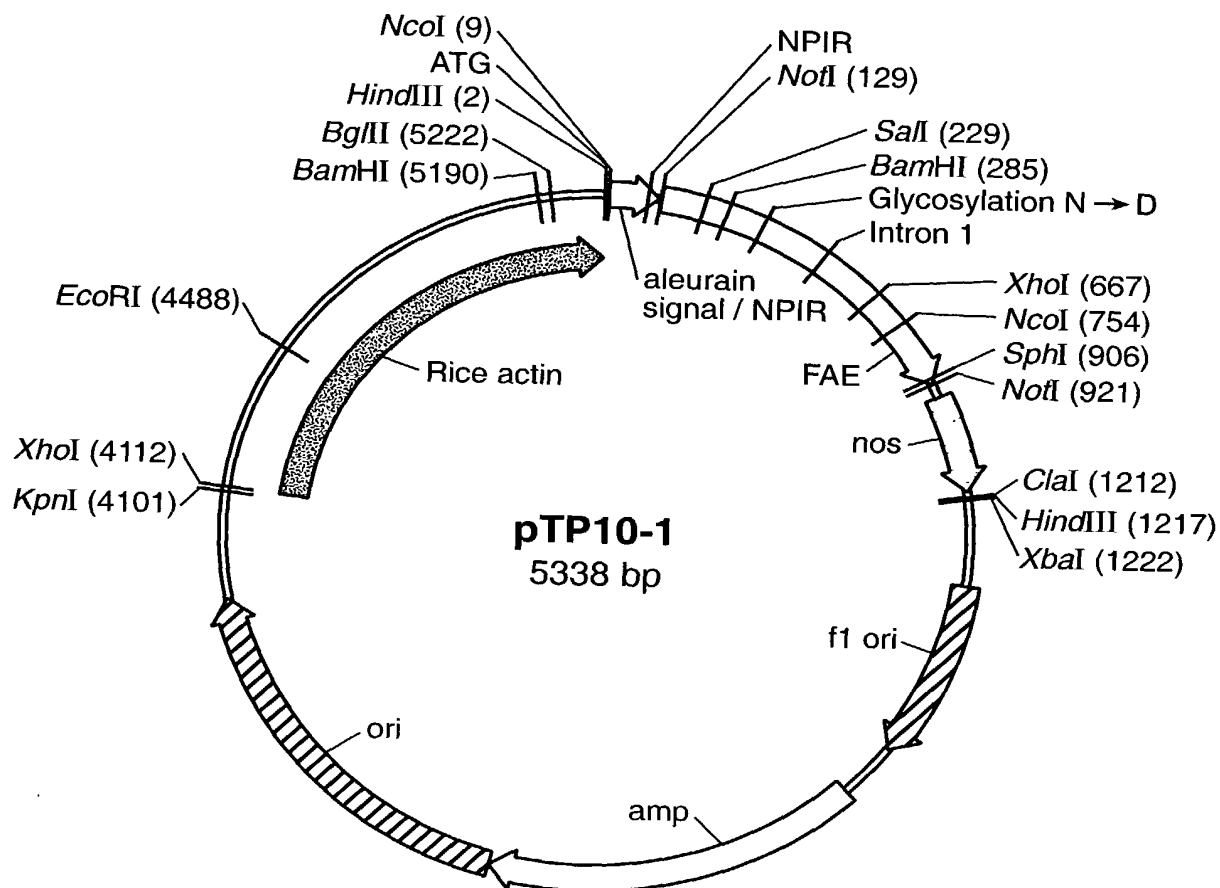


FIG._32A

36 / 154

```

      NcoI
      ~~~~~
HindIII
~~~~~
      M A H A R V L L L A L A V L A T A A V A V
1 AAGCTTACCA TGGCCACGC CCGGTCTCTC CTCCTGGCGC TCGCCGTGCT GGCACGGCC GCCGTGCGC

      NPIR
      ~~~~~
      . A S S S F A D S N P I R P V T D R A A A S T .
71 TCGCCTCCTC CTCCTCCTTC GCGACTCCA ACCGATCCG GCCGTCACC GACCGCGGG CCGCCTCCAC
  . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCAATA TCTCCCAAGC TGCCTACGCC

      NotI
      ~~~~~
      . A S S S F A D S N P I R P V T D R A A A S T .
71 TCGCCTCCTC CTCCTCCTTC GCGACTCCA ACCGATCCG GCCGTCACC GACCGCGGG CCGCCTCCAC
  . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCAATA TCTCCCAAGC TGCCTACGCC

      Sali
      ~~~~~
      D L C N I P S T I I K G E K I Y N S Q T D I N G
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGAGAGA AAATTTACAA TTCTCAAACT GACATTAACG

      BamHI
      ~~~~~
      . W I L R D D S S K E I I T V F R G T G S D T N .
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACCTGGTA GTGATACGAA

      Glycosylation
      ~~~~~
      . L Q L D T D Y T L T P F D T L P Q C N G C E V
351 TCTACAACTC GATACTGACT ACACCTTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
      H G G Y Y I G W V S V Q D Q V E S L V K Q Q V S
421 CACGGTGGAT ATTATATTGG ATGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAA CAGCAGGTTA
  . Q Y P D Y A L T V T G H X L G A S L A A L T A .
491 GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
  . A Q L S A T Y D N I R L Y T F G E P R S G N Q
561 CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
  
```

FIG..32B

37 / 154

```

XhoI
~~~~~
631  A F A S Y M N D A F Q A S S P D T T Q Y F R V T
    GCCTTCGCGT CGTACATGAA CGATGCCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGGTCA

NcoI
~~~~~
701  . H A N D G I P N L P P V E Q G Y A H G G V E Y .
    CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
    . W S V D P Y S A Q N T F V C T G D E V Q C C E
    CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTGTC TGCACCTGGG ATGAAGTGCA GTGCTGTGAG

SphI
~~~~~
841  A Q G G Q G V N N A H T T Y F G M T S G A C T W
    GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGC GCATGCACCT

NotI
~~~~~
KDEL
~~~~~
911  . P V A A A E T T E G *
    GGCCGGTCGC GGCCGGCGAA ACCACTGAAG GATGAGCTGT AAAGAAGCAG ATCGTTCAA CATTGGCAA
981  TAAAGTTTCT TAAGATTGAA TCCTGTGTC GGTCTGCGA TGATTATCAT ATAATTTCTG TTGAATTACG
1051 TTAAGCATGT AATAATTAA ATGTAATGCA TGACGTTATT TATGAGATGG GTTTTATGA TTAGAGTCCC
1121 GCAATTATAC ATTTAATACG CGATAGAAA CAAAATATAG CGCGCAAACT AGGATAAATT ATCGCGCGCG

HindIII
~~~~~
ClaI
~~~~~
XbaI
~~~~~
1191 GTGTCATCTA TGTTACTAGA TCGATAAGCT TCTAGAGCGG CCGGTGGAGC TCCAATTGCG CCTATAGTGA
1261 GTCGTATTAC GCGCGCTCAC TGGCCGTCGT TTACAAACGT CGTGACTGGG AAAACCCCTGG CGTTACCCAA
1331 CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC
1401 CTTCCCAACA GTTGGCGAGC CTGAATGGCG AATGGGACGC GCCCTGTAGC GGCGCATTA GCGCGGCGGG
```

FIG.-32C

1471 TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTCTTTC
1541 CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCCTCAAG CTCTAAATCG GGGGCTCCCT TTAGGGTTCC
1611 GATTAGTGC TTACGGCAC CTCGACCCCA AAAAATTGA TTAGGGTGAT GGTTCACGTA GTGGGCCATC
1681 GCCCTGATAG ACGGTTTTTC CCGCTTTGAC GCTTCTTTG ATTTATAAGG GATTTTGCCG ATTTGCGCCT
1751 ACTGGAACAA CACTCAACCC CACTCAACCC TATCTCGGTC TATCTTTTG ATTTATAAGG GATTTTGCCG ATTTGCGCCT
1821 ATTGGTTAAA AAATGAGCTG AATTAACAAA AATTTAACGC GAATTTTAA TTTCTTAAATA CATTCAAAAT
1891 TTAGGTGGCA CTTTTCGGG CATGAGACAA TAACCTTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
1961 TGTATCCGCT CATGAGACAA TAACCTTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
2031 TCAACATTTT CGTGTGCGCC TATTTCCCTT TGCTGAAGAT CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA
2101 ACGCTGGTGA AAGTAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA
2171 ACAGCGGTAA GATCCTTGAG GATCCTTGAG AGTTTTCGCC CCGAAGAACG TTTTCCCAATG ATGAGCACTT TTAAGTTCT
2241 GCTATGTGCG CCGGTATTAT CCGGTATTGA CTCACCAATG ACAGTGGGTG TTTTCCCAATG ATGAGCACTT TTAAGTTCT
2311 CAGAAATGACT TGGTTGAGTA CTCACCAATG ACAGTGGGTG TTTTCCCAATG ATGAGCACTT TTAAGTTCT
2381 TATGCAGTGC TGCCATAACC ACCGCTTTT TGCACCAATG ACAGTGGGTG TTTTCCCAATG ATGAGCACTT TTAAGTTCT
2451 GAAGGAGCTA ACCGCTTTT TGCACCAATG ACAGTGGGTG TTTTCCCAATG ATGAGCACTT TTAAGTTCT
2521 CTGAATGAAG CCATACCAAA TGGCGAACTA CTTACTCTAG GCTCGGCCCT TCCGGCTGGC ACAAATTAATA GACTGGATGG AGCGGATAA
2591 AACTATTAACT TGGCGAACTA CTTACTCTAG GCTCGGCCCT TCCGGCTGGC ACAAATTAATA GACTGGATGG AGCGGATAA
2661 AGTTGCAGGA CCATTTCTGC CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT
2731 GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT
2801 ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
2871 TAAGCATGG TAACGTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATG ATTTAAAACT TCATTTTAA
2941 TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT CCGTTAAACGT GAGTTTTCGT
3011 TCCACTGAGC GTCAGACCC GTAGAAAAGA TCAAGGATC TTTCTGAGAT CTTTGTGATG TCGCGGTAAT
3081 CTGCTGCTTG CAAACAAAAG AGGTAACCTG CTTTCAAGAAC TCTGTAGCAC CGCTTACATA CTTTGTGATG TCGCGGTAAT
3151 CTTTTCCTGA TAGGCCACCA CTTTCAAGAAC TCTGTAGCAC CGCTTACATA CTTTGTGATG TCGCGGTAAT
3221 TAGGCCACCA CTTTCAAGAAC TCTGTAGCAC CGCTTACATA CTTTGTGATG TCGCGGTAAT
3291 TGCTGCCAGT GCGGATAAGT GCGGATAAGT TCTGTGACCA CAGCCCACTG TGGAGCGAAC GACCTACACC GAACTGAGAT
3361 CCGTCGGGCT GAACGGGGG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGAGAAAG CCGGACAGGT ATCCGGTAAG
3431 ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGAGAAAG CCGGACAGGT ATCCGGTAAG
3501 CGGCAGGGTC GGAACAGGAG AGCGCACGAG GAGCTTTTGT GATGCTCGTC AGGGGGCGG CCTGGTATCT TTATAGTCTT
3571 GTCGGGTTTC GCCACCTCTG ACTTGAGCGT TTTTACCGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTCTTTTC
3641 AAAACGCCAG CAACGCCGCT CCGGATAAGT TTTTACCGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTCTTTTC
3711 TGCCTTATCC CCGGATAAGT TTTTACCGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTCTTTTC
3781 CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCC AATACGCAAA CCGCCTCTCC
3851 CCGCGCGTTG GCCGATTCTA TAATGCAGCT GGCACGACAG GTTTCCCGAC TTTTCCCGAC TGGAAAGCGG GCAGTGAGCG

FIG.-32D

39 / 154

```
3921 CAACGCAATT AATGTGAGTT AGTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT
3991 ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG TTTACACGCTA TGACCATGAT TACGCCAAGC

                                KpnI
                                ~~~~~
4061 GCGCAATTAA CCTCACTAA AGGGAACAAA AGTGGGTAC CGGGCCCCCC TTCGAGGTCA TTCATATGCT
4131 TGAGAAGAGA GTCGGGATAG TCCAAAATAA AACAAAGGTA AGATTACCTG GTCAAAAGTG AAAACATCAG
4201 TTAAAAGGTG GTATAAGTAA AATATCGGTA ATAAAAGGTG GCCCAAAGTG AAATTTACTC TTTTCTACTA
4271 TTATAAAAAAT TGAGGATGTT TTGTCGGTAC TTTGATACGT CATTTTGTGA TGAATTGGTT TTTAAGTTTA
4341 TTCGCGATTT GGAAATGCAT ATCTGTATTT GAGTCGGTTT TTAAGTTTCGT TGCTTTTGTG AATACAGAGG
4411 GATTGTGATA AGAAATATCT TTAAAAAACC CATATGCTAA TTTGACATAA TTTTGTGAGAA AAATATATAT

                                EcoRI
                                ~~~~~
4481 TCAGGCGAAT TCCACAATGA ACAATAATA GATTAAATA GCTTGCCCCC GTTGCAGCGA TGGGTATTTT
4551 TTCTAGTAAA ATAAAAGATA AACTTAGACT CAAAACATTT ACAAAACAA CCCCTAAAAGT CCTAAAGCCC
4621 AAAGTGCTAT GCACGATCCA TAGCAAGCCC AGCCCAACCC AACCCACCCC AGCCCAACCC AGTGCAGCCA
4691 ACTGGCAAT AGTCTCCACC CCGGCACTA TCACCGTGAG TTGTCGGTCC CACCGCACGT CTCGCAGCCA
4761 AAAAAAAA AAGAAAGAAA AAAAAAGAAA AAAAAAGAAA CAGGTGGTCC CGGTCGTGG GGGCCGAAA
4831 AGCGAGGAGG ATCGCGAGCA GCGACGAGG CCGGCCCTCC CTCCGCTTCC AAAGAAACGC CCCCCTCGC
4901 CACTATATAC ATACCCCCC CTCTCCTCCC ATCCCCCAA CCTACCACC ACCACCACCA CCACCTCCTC
4971 CCCCCTCGCT GCCGGACGAC GAGCTCCTCC CCCCCTCCCC TCCGCCGCGG CCGGTAACCA CCCCGCCCTT
5041 CTCCTCTTTC TTCTCCGTT TTTTTTTTCG TCTCGGTCTC GATCTTTGGC CTTGGTAGTT TGGGTGGCG
5111 AGAGCGGCTT CGTCGCCAG ATCGGTGCGC GGGAGGGCG GGATCTCGCG GCTGGCGTCT CCGGGCGTGA

                                BamHI
                                ~~~~~
5181 GTCGGCCCCG ATCCTCGCG GGAATGGGC TCTCGGATGT AGATCTTCTT TCTTCTTCT TTTTGTGTA
5251 GAATTGAAAT CCTCAGCAT TGTTTCATCG TAGTTTTTCT TTTTCATGAT TGTGACAAAT GCAGCCTCGT
5321 GCGGAGCTTT TTTGTAGC
```

FIG._32E

40 / 154

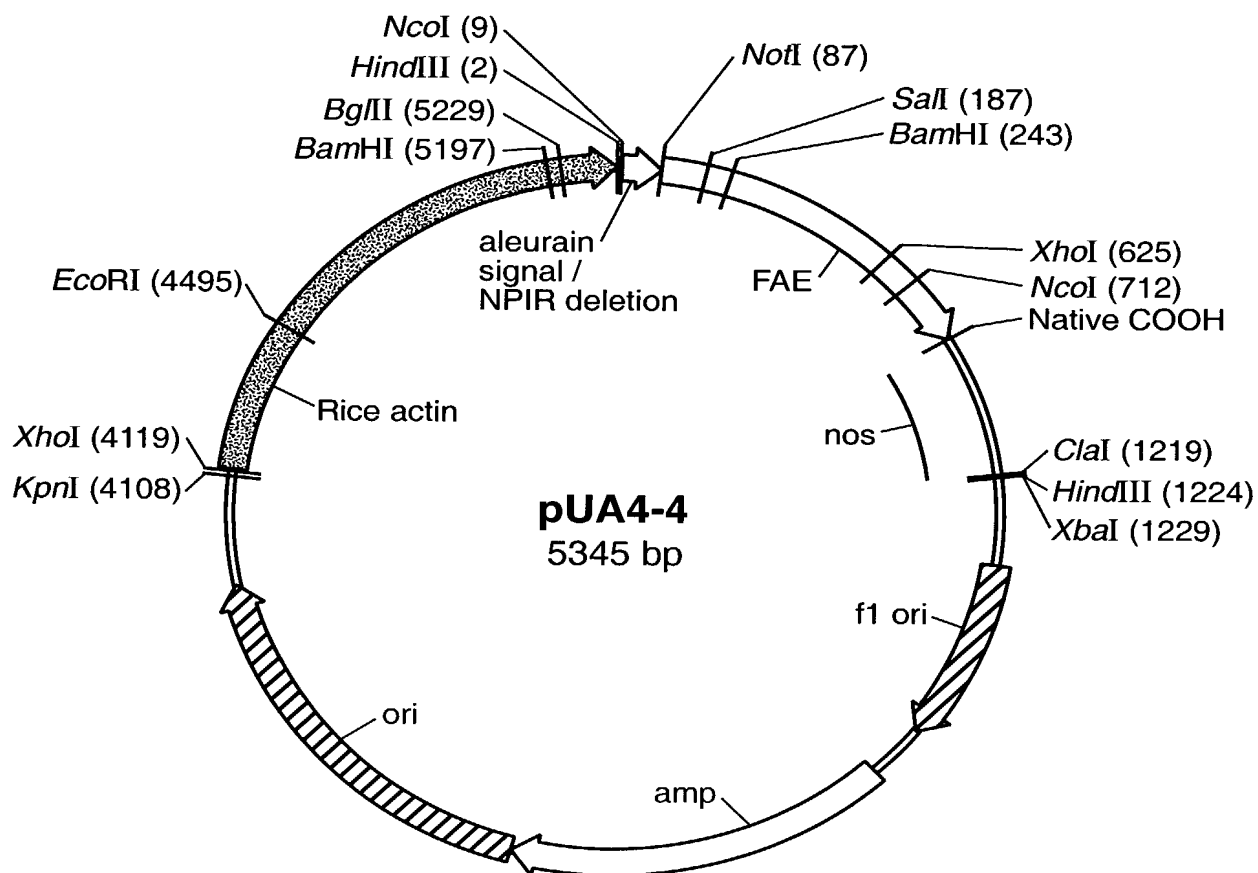


FIG._33A

41 / 154

```

NcoI
~~~~~
HindIII
~~~~~
      M A H A R V L L L A L A V L A T A A V A V
1  AAGCTTACCA TGGCCACAGC CCGGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG

      . A S S R A A A S T Q G I S E D L Y S R L V E M .
71 TCGCCTCCTC CCGCGCGGCC GCCTCCACGC AGGCATCTC CGAAGACCTC TACAGCCGTT TAGTCGAAAT
      Sali
      ~~~~~
      . A T I S Q A A Y A D L C N I P S T I I K G E K
141 GGCCACTATC TCCAAGCTG CCTACGCCGA CTTGTGCAAC ATTCCGTCGA CTATTATCAA GGGAGAGAAA
      BamHI
      ~~~~~
      I Y N S Q T D I N G W I L R D D S S K E I I T V
211 ATTTACAATT CTCAAACTGA CATTACGGA TGGATCCTCC GCGACGACAG CAGCAAAGAA ATAATCACCG
      . F R G T G S D T N L Q L D T N Y T L T P F D T .
281 TCTTCCGTGG CACTGGTAGT GATACGAATC TACAACCTCGA TACTAACTAC ACCCTCACGC CTTTCGACAC
      . L P Q C N G C E V H G G Y Y I G W V S V Q D Q
351 CCTACCACAA TGCAACGGTT GTGAAGTACA CGGTGGATAT TATATTGGAT GGTCTCCGT CCAGGACCAC
      V E S L V K Q Q V S Q Y P D Y A L T V T G H X L
421 GTCGAGTCGC TTGTCAAACA GCAGGTAGC CAGTATCCGG ACTACGGCT GACCGTGACC GGCCACKCCC
      . G A S L A A L T A A Q L S A T Y D N I R L Y T .
491 TCGGCGCCTC CCTGGCGGCA CTCACTGCCG CCCAGCTGTC TCGGACATAC GACAACATCC GCCTGTACAC
      XhoI
      ~~~~~
      . F G E P R S G N Q A F A S Y M N D A F Q A S S
561 CTTGGCGGAA CCGCGCAGCG GCAATCAGCG CTTGCGGTG TACATGAACG ATGCTTCCA AGCCTCGAGC
      P D T T Q Y F R V T H A N D G I P N L P P V E Q
631 CCAGATACGA CGCAGTATTT CCGGGTCACT CATGCCAAGC ACGGCATCCC AAACCTGCC CCGGTGGAGC
      NcoI
      ~~~~~
      . G Y A H G G V E Y W S V D P Y S A Q N T F V C .
```

FIG._33B

```
701 AGGGGTACGC CCATGGCGGT GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCCAGAACA CATTGTGCTG
    . T G D E V Q C C E A Q G G Q G V N A H T T Y
771 CACTGGGGAT GAAGTGCAGT GCTGTGAGGC CCAGGGCGGA CAGGTTGTGA ATAAATGCGCA CACGACTTAT
    F G M T S G A C T W *
841 TTTGGGATGA CGAGCGGAGC CTGTACATGG TGATCAGTCA TTTCAGCCTC CCCGAGTGTA CCAGGAAAGA
911 TGGATGTCCT GGAGAGGGGG CCGCGTAACC ACTGAAGGAT GAGCTGTAAA GAAGCAGATC GTTCAAACAT
981 TTGGCAATAA AGTTTCTTAA GATTGAATCC TGTGCGCGT CTTGCGATGA TTATCATATA ATTTCTGTTG
1051 AATTACGTTA AGCATGTAAT AATTAACATG TAATGCATGA CGTTATTTAT GAGATGGGTT TTTATGATTA
1121 GAGTCCCGCA ATTATACATT TAATACGCGA TAGAAAACAA AATATAGCGC GCAAACCTAGG ATAAATTATC
    HindIII
    ~~~~~
    ClaI XbaI
    ~~~~~
1191 GCGCGCGGTG TCATCTATGT TACTAGATCG ATAAAGCTTCT AGAGCGGCGG GTGGAGCTCC AATTCGCCCT
1261 ATAGTGAGTC GTATTACGGC CGCTCAGTGG CCGTCGTTTT ACAACGTCGT GACTGGGAAA ACCCTGGCGT
1331 TACCCAACTT AATCGCCCTG CAGCACATCC CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCTGCACC
1401 GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGGACGCGCC CTGTAGCGGC GCATTAAGCG
1471 CGGCGGGTGT GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGGCCG CTCTTTTCGC
1541 TTTCTTCCCT TCCTTTCTCG CCACGTTCCG CGGCTTTCCC CGTCAAGCTC TAAATCGGGG GCTCCCTTTA
1611 GGGTTCGGAT TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTA GGGTGATGGT TCACGTAGTG
1681 GGCCATCGCC CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT
1751 GTTCCAAACT GGTAAATAA TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAAACAA ATATTAACGC
1821 TCGGCCCTAT TACAATTTA GGTGGCACTT TTCGGGAAA CCTGATAAA TGCTTCAATA ATATTGAAA AGGAAGAGTA
1891 TCAAATATGT ATCCGCTCAT GAGACAATAA CCGTCCCTTA TTCCCTTTT ACCCTATTT GTTTATTTT CTAAATACAT
1961 TGAGTATTCA ACATTCCGT GTCCGCCCTA TTCCCTTTT TGCGGCATTT TGCCCTTCCCTG TTTTGTGCTCA
2031 CCCAGAAAAC CTGGTGAAAG TAAAGATGCG TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAAGTG
2101 GATCTCAACA GCGGTAAAGT CCTTGAGAGT TTTCCGCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA
2171 AAGTCTGCT ATGTGGCGG GTATTATCCC GTATTGACG CCGGCAAGAG CAACCTCGGC GCCGCATACA
2241 CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAGCATC TTACGGATGG CATGACAGTA
2311 AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG
2381 GAGGACCGAA GGAGCTAACC GCTTTTTCG ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA
2451 ACCGGAGCTG AATGAAGCCA TACCAAAACG CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG
2521 TTGCGCAAAC TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG
2591 CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG
2661
```

FIG._33C

```
2731 AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA
2801 GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT
2871 CACTGATTAA GCATTGGTAA CTGTGAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACCTTCA
2941 TTTTAAATTT AAAAGGATCT AGTGAAGAT CTTTTTGAT AATCTCATGA CCAAAATCCC TTAAACGTGAG
3011 TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAAGATCA AAGGATCTTC TGTGAGATCCT TTTTCTCTGC
3081 GCGTAATCTG CTGCTTGCAA ACAAAAAAC CACCGCTACC AGCGGTGGTT AGCGGTCCCG ATCAAGAGCT
3151 ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA ATACTGTCTT TCTAGTGTAG
3221 CCGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC
3291 CAGTGGCTGC TGCAGTGGC GATAAGTCGT GTCTTACC GGTTGGACTCA AGACGATAGT TACC GGATAA
3361 GCGCAGCGG TCGGGCTGAA CCGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA
3431 CTGAGATACC TACAGCGTGA GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGCGG GACAGGTATC
3501 CGGTAAGCGG CAGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAACGCCCT GGTATCTTTA
3571 TAGTCCCTGC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTGTGTGAT GCTCGTCAGG GGGCGGGAGC
3641 CTATGGAATA ACGCCAGCAA CGCGGCCCTT TTACGGTTCC TGGCCCTTTT CTGGCTTTT GCTCACATGT
3711 TCTTCTCTGC GTTATCCCTT GATTCTGTGG ATAAACCGTAT TACCGCCCTT GAGTGAAGTG ATACCGCTCG
3781 CCGCAGCCGA ACGACCGAGC GCAGCGAGTC AGTGAGCGAG GAAAGCGGAG AGCGCCCAAT ACGCAAAACCG
3851 CCTCTCCCGG CGCGTTGGCC GATTCAATTA TGCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGGCA
3921 GTGAGCGCAA CGCAATTAAT GTGAGTTAGC TCACTCATTA GGCACCCCGAG GCTTTACACT TTATGCTTCC
3991 GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCAATGATTAC

                                     KpnI
                                     ~~~~~
4061 GCCAAGCGCG CAATTAAACC TCACTAAAGG GAACAAAAGC TGGGTACCGG GCCCCCCCTC GAGGTCATTC
4131 ATATGCTTGA GAAGAGAGTC GGGATAGTCC AAAATAAAAC AAAGGTAAGA TTACCTGGTC AAAAGTGAAA
4201 ACATCAGTTA AAAGGTGGTA TAAGTAAAT ATCGGTAATA AAAGGTGGCC CAAAGTGAAA TTACTCTTTT
4271 TCTACTATTA TAAAAATTGA GGATGTTTTG TCGGTACTTT GATACGTCAT TTTTGTATGA ATTGGTTTTT
4341 AAGTTTATTC GCGATTGGA AATGCATATC TGTATTTGAG TCGGTTTTTA AGTTCGTTGC TTTTGTAAAT
4411 ACAGAGGGAT TTGTATAAGA AATATCTTTA AAAAAACCAT ATGCTAATTT GACATAATTT TTGAGAAAAA

                                     EcoRI
                                     ~~~~~
4481 TATATATTCA GGCGAATTCC ACAATGAACA ATAATAAGAT TAAAATAGCT TGCCCCCGTT GCAGCGATGG
4551 GTATTTTTTC TAGTAAAAATA AAAGATAAAC TTAGACTCAA AACATTTACA AAAACAACCC CTAAAGTCCT
4621 AAAGCCCAAA GTGCTATGCA CGATCCATAG CAAGCCCAAGC CCAACCCAAC CCAACCCAGT
4691 GCAGCCAACT GGCAAAATAGT CTCACCCGCC GGCACATATCA CCGTGAGTTG TCCGCAACAC CGCACGCTC
4761 GCAGCCAAAA AAAAAAAAG AAGAAAAAA AAAAAAAGA AAAACAGCAG GTGGGTCCCG GTCGTGGGGG
4831 CCGGAAAAGC GAGGAGGATC GCGAGCAGCG ACGAGGCCCG GCCCTCCCTC CGCTTCCAAA GAAACGCCCC
```

FIG._33D

44 / 154

```
4901 CCATCGCCAC TATATACATA CCCCCCCCCTC TCCTCCCATC CCCCACAACC TACCACCACC ACCACCACCA
4971 CCTCCTCCCC CCTCGCTGCC GGACGACGAG CTCCTCCCCC CTCCCCCTCC GCCGCCGCCG GTAACCAACC
5041 CGCCCCCTCTC CTCCTTCTTT CTCCGTTTTT TTTTTCGTCT CCGTCTCGAT CTTTGGCCCTT GGTAGTTTGG
5111 GTGGGCGAGA GCGGCTTCGT CGCCAGATC CGCCGCGGG GGTGCGCGG AGGGCGGGA TCTCGCGGCT GCGCTCTCCG
                                     BamHI
                                     ~~~~~
5181 GCGGTGAGTC GGCCCGGATC CTCGCGGGGA ATGGGGCTCT CGGATGTAGA TCTTCTTTCT TTCTTCTTTT
5251 TGTGGTAGAA TTTGAATCCC TCAGCATTGT TCATCGGTAG TTTTCTTTTT CATGATTGT GACAAATGCA
5321 GCCTCGTGCG GAGCTTTTTT GTAGC
```

FIG._33E

45 / 154

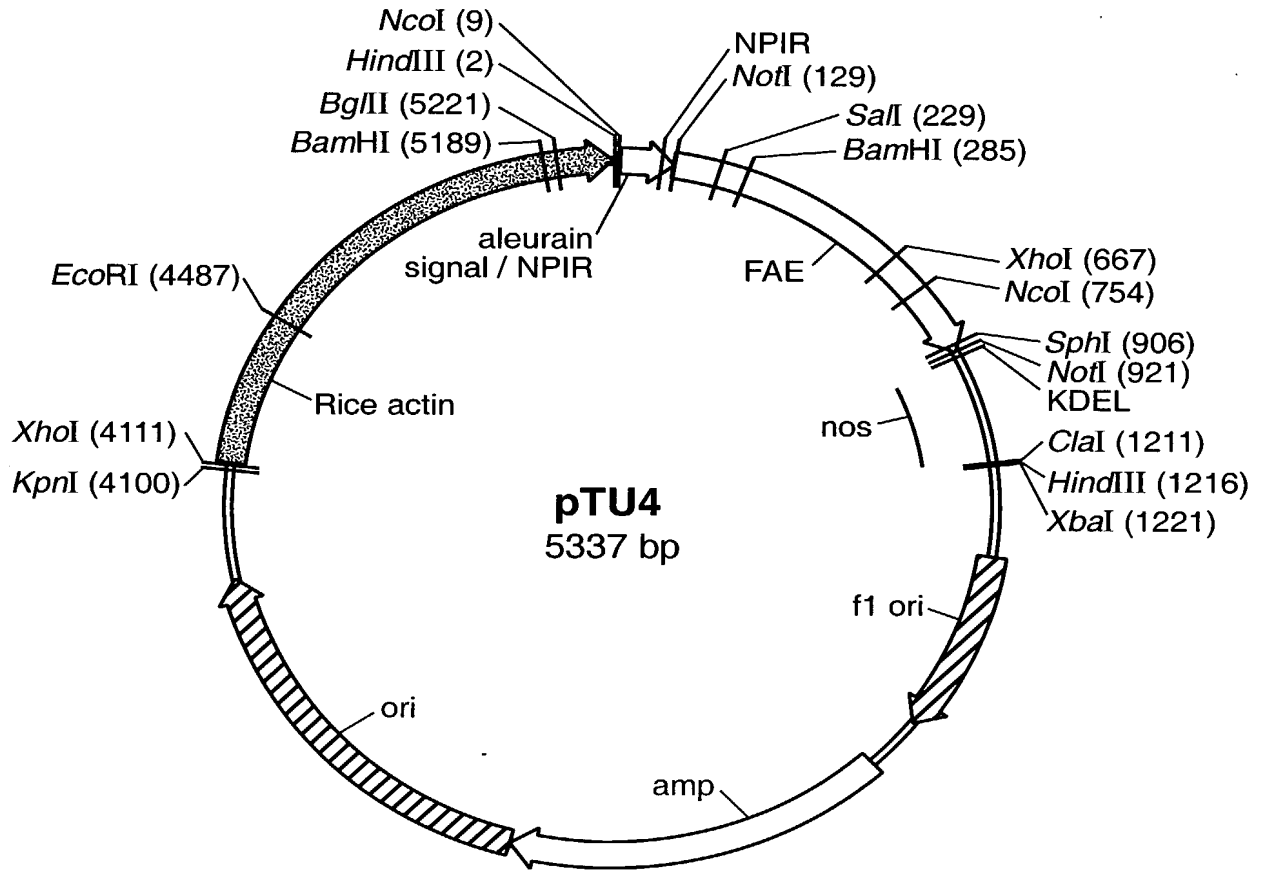


FIG._34A

46 / 154

```

NcoI
~~~~~
HindIII
~~~~~
      M A H A R V L L L A L A V L A T A A V A V
1 AAGCTTACCA TGGCCCACGC CCGCGTCCCTC CTCCTGGCGC TCGCCGTGCT GCCACGGCC GCCGTCGCCG
      NotI

      . A S S S F A D S N P I R P V T D R A A A S T .
71 TCGCCTCCTC CTCCTCCTTC GCGACTCCA ACCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
      . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC
      SalI
      ~~~~~
      D L C N I P S T I I K G E K I Y N S Q T D I N G
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGAGAGA AAATTACAA TTCTCAAACT GACATTAACG
      BamHI
      ~~~~~
      . W I L R D D S S K E I I T V F R G T G S D T N .
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACCTGGTA GTGATACGAA
      . L Q L D T N Y T L T P F D T L P Q C N G C E V
351 TCTACAACTC GATACTAACT ACACCTCAC GCCTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
      H G G Y Y I G W V S V Q D Q V E S L V K Q Q V S
421 CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAA CAGCAGGTTA
      . Q Y P D Y A L T V T G H X L G A S L A A L T A .
491 GCCAGTATCC GGAATACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
      . A Q L S A T Y D N I R L Y T F G E P R S G N Q
561 CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
      XhoI
      ~~~~~
      A F A S Y M N D A F Q A S S P D T T Q Y F R V T
631 GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCAGATAC GACGCAGTAT TTCCGGGTCA
      NcoI
      ~~~~~
      . H A N D G I P N L P P V E Q G Y A H G G V E Y .

```

FIG.-34B

47 / 154

```
701 CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCCTGGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
    . W S V D P Y S A Q N T F V C T G D E V Q C C E
771 CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTGTC TGCACCTGGG ATGAAGTGCA GTGCTGTGAG
    SphI
    ~~~~
    A Q G G Q G V N N A H T T Y F G M T S G A C T W
841 GCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGC GCATGCACCT

    NotI
    ~~~~~~
    . P V A A A E P L K D E L *
911 GGCCGGTCGC GGCCGGGAA CCACCTGAAGG ATGAGCTGTA AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT
981 AAAGTTTCTT AAGATTGAAT CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTTCTGT TGAATTACGT
1051 TAAGCATGTA ATAATTAAACA TGAATGCAT GACGTTATTT ATGAGATGGG TTTTATGAT TAGAGTCCCG
1121 CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG

    HindIII
    ~~~~~~
    ClaI XbaI
    ~~~~~~
1191 TGTCATCTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC CTATAGTGAG
1261 TCGTATTACG CGCGCTCACT GGCCGTCTGT TTACAACGTC GTGACTGGGA AAACCTTGGC GTTACCCCAAC
1331 TTAATCGCCT TGCAGCACAT CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC
1401 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGACGCG CCCTGTAGCG GCGCATTAAG CGCGCGGGGT
1471 GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTCTTCC
1541 CTTCCCTTCT CGCCACGTTT TACCGGCACC TCGACCCCAA AAACTTGAT TAGGTCCCTT TAGGGTTCCG
1611 ATTTAGTGCT TTACGGCACC CCCTTTTTCG CCCTTTGACG TTGGAGTCCA CGTCTTTAA TAGTGGACTC
1681 CCCTGATAGA CGGTTTTTCG ATCTCGGTCT ATCTTTTGA TTTATAAGG ATTTTGCCGA TTTTCGGCCTA
1751 CTGGAACAAC ACTCAACCTT AATGAGCTGA TTTAACAAA AATTTAACG AATTTAACG GCTTACAATT
1821 TTGGTTAAA AATGAGCTGA AATGCGGGA AATGCGGGA AATGCTTCA TAATATTGAA AAAGGAAGAG
1891 TAGGTGGCAC TTTTCGGGCA ATGAGACAAT AACCTGATA AATGCTTCA TAATATTGAA AAAGGAAGAG
1961 GTATCCGCTC GTGTCGCCCT TATTCCTTTT TTTGCGGCAT TTTGCTTCC ACAGTGGGT TACATCGAAC
2031 CAACATTTC CGCTGGTAA AGTAAAGAT GCTGAAGATC AGTTGGGTGC ACAGTGGGT TTTTCCAAATGA
2101 CAGCGGTAAG ATCCTTGAGA GTTTTCGCC CCGAAGACGT TTTTCCAAATGA ACAGTGGGT TTTTCCAAATGA
2171 CTATGTGGCG CGGTATTATC CCGTATTGAC GCGGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
2241 AGAATGACTT GGTGAGTAC TCACCACTCA CAGAAAAGCA TCTTACGGAT GGCAATGACAG TAAAGAGAATT
2311 ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACCGAT CGGAGGACCG
2381
```

FIG._34C

48 / 154

```
2451 AAGGAGCTAA CCGCTTTTTC GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC
2521 TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA
2591 ACTATTAACT GCGGAACCTAC TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GCGGGATAAA
2661 GTTGACAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTATTATTC TGGTAAAGCCC TAGTTATCTA
2731 AGCGTGGGTC TCGCGGTATC ATGTCAGCAC CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT
2801 CACGACGGGG AGTCAGGCAA AACTGTCTAG CCAAGTTTAC TCATATATAC TTTAGATTGA TTTTAAACTT CATTTTAAAT
2871 AAGCATTTGGT AACTGTCTAG CCAAGTTTAC TCATATATAC TTTAGATTGA TTTTAAACTT CATTTTAAAT
2941 TTTAAAGGAT CTAGGTGAAG ATCCTTTTTC TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTCTT GCGCGTAATC
3011 CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTCTT GCGCGTAATC
3081 TGCTGCTTGC AAACAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC
3151 TTTTTCGGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT AGCGTAGTGT
3221 AGCCACCCAC TTCAAGAACT CTGTAGCACC GCTTACATAC CTCGCTCTGC TAAATCCTGTT ACCAGTGGCT
3291 GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGTGTGGACT CAAGACGATA GTTACCCTGAT AAGGCGCAGC
3361 GGTGCGGCTG AACGGGGGGT TCGTGACAC ACCTCAGCTT GGAGCGAAGG ACCTACACCG AACTGAGATA
3431 CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CCGACAGGTA TCCGGTAAGC
3501 GGCAGGGTCG GAACAGGAGA GCGCAGGAG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCCTG
3571 TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTGTG ATGCTCGTCA GGGGGCGGA GCCTATGGAA
3641 AAACGCCAGC AACCGGGCTT TTTTACGGTT CTTGGCCCTT TGCTGGCCTT TTGCTCACAT GTTCTTTCTT
3711 GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCTT TTGAGTGAGC TGATACCGCT CGCGCAGCC
3781 GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCA ATACGCAAAC CGCTCTCCC
3851 CGCGCGTTGG CCGATTCAAT AATGCAGCTG GCACGACAGG TTTCCCGACT GGAAGCGGG CAGTGAGCGC
3921 AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA
3991 TGTTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG AACACAGCTAT GACCATGATT ACGCCAAGCG

4061 CGCAATTAA C CTTCACTAAA GGAACAAAA GCTGGGTACC GGGCCCCCCC TCAGAGGTCT TCATATGCTT
4131 GAGAAGAGAG TCGGGATAGT CCAAAATAAA ACAAAAGTAA GATTACCTGG TCAAAAAGTGA AAACATCAGT
4201 TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAAGGTGG CCAAAAGTGA AATTTACTCT TTTCTACTAT
4271 TATAAAAAAT GAGGATGTTT TGTCGGTACT TTGATACGTC ATTTTGTAT GAATTGGTTT TTAAGTTTAT
4341 TCGCGATTGG GAAATGCATA TCTGTATTGG AGTCGGTTT TAAAGTTCGTT GCTTTGTAA ATACAGAGGG
4411 ATTTGTATAA GAAATATCTT TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA AATATATATT

      ECORI
      ~~~~~
4481 CAGGCGAATT CCACAATGAA CAATAATAAG ATTAAAAATAG CTTGCCCCCC TTGCAGCGAT GGGTATTTTT
4551 TCTAGTAAAA TAAAAGATAA ACTTAGACTC AAAACATTTA CAAAAACAAC CCTTAAAGTC CTAAAGCCCA
```

FIG.-34D

09/09/2009 10:20:00

49 / 154

```
4621 AAGTGCTATG CACGATCCAT AGCAAGCCCA GCCAAACCCA ACCAAACCCA ACCACCCCA GTGCAGCCAA
4691 CTGGCAAATA GTCTCCACCC CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC TCGCAGCCAA
4761 AAAAAAAAAA AGAAGAAAAA AAAAGAAAAA GAAAAACAGC AGGTGGGTCC GGTTCGTGGG GGCCGGAAAA
4831 GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC CCCCATCGCC
4901 ACTATATACA TACCCCCCCC TCTCCTCCCA TCCCCCAAC CTTACCACCA CCACCACCAC CACCTCCTCC
4971 CCCCTCGCTG CCGGACGACG AGCTCCTCCC CCCTCCCCCT CCGCCGCCGC CGGTAACCCAC CCCGCCCTC
5041 TCCTCTTTCT TTCTCCGTTT TTTTCTTCTG CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT GGGTGGGGCA
5111 GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC CGGGCGTGAG

                                     BamHI
                                     ~~~~~
5181 TCGGCCCCGA TCCTCGCGGG GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG
5251 AATTGAAATC CCTCAGCAAT GTTCATCGGT AGTTTTTCTT TTCAATGATTT GTGACAAATG CAGCCTCGTG
5321 CGGAGCTTTT TTGTAGC
```

FIG._34E

50 / 154

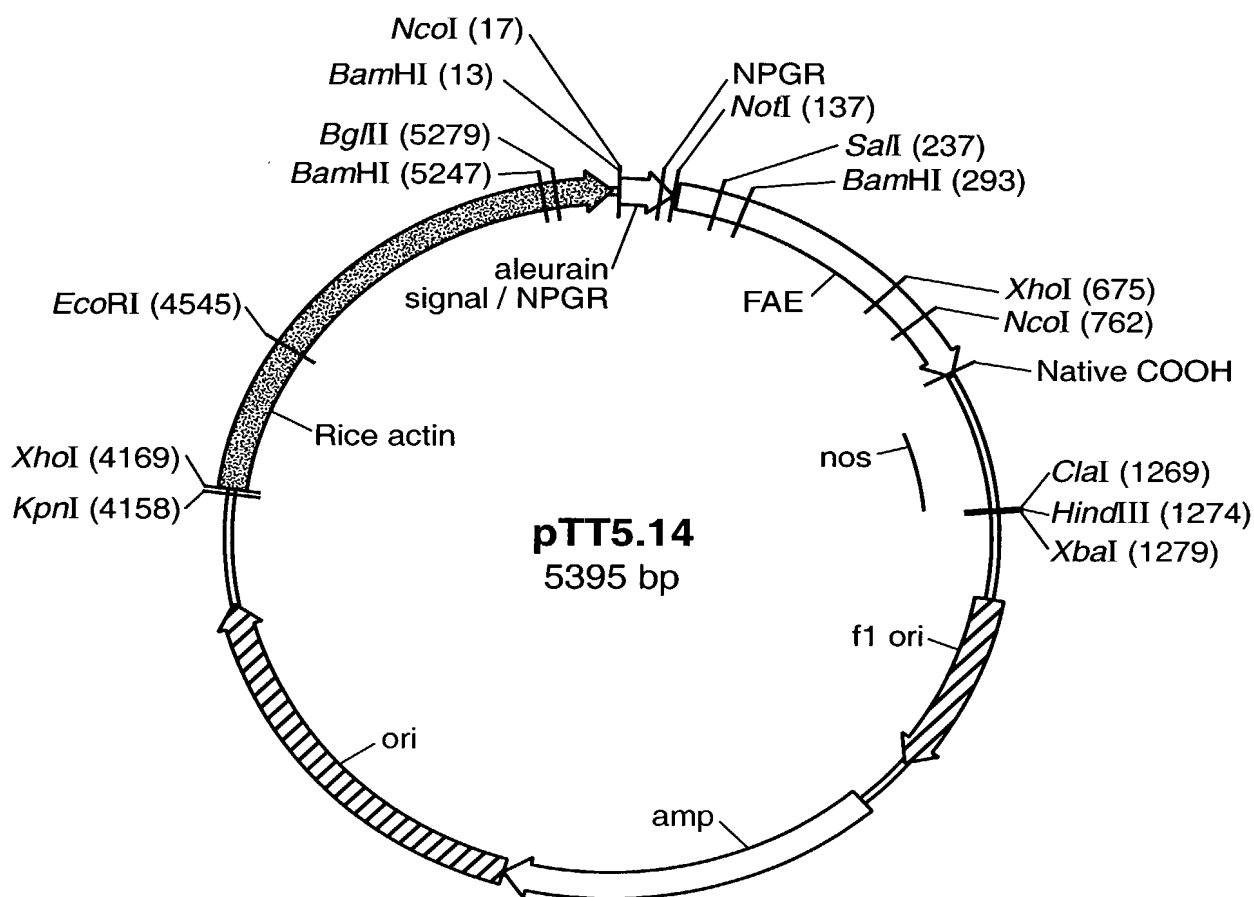


FIG._35A

51 / 154

```

NcoI
~~~~~
BamHI
~~~~~
1  CCTGACGCCG AGGATCCATG GCCACGCC GGTCTCTCCT CTTGGCGCTC GCCGTGCTGG CCACGGCGCG
    M A H A R V L L L A L A V L A T A A .
    NotI
    . V A V A S S S F A D S N P G R P V T D R A A
71  CGTCGCCGTC GCCTCTCTCT CCTCTCTCGC CGACTCCAAC CCGGGCCGGC CCGTCACCGA CCGCGCGGGC
    NotI
    .
    A S T Q G I S E D L Y S R L V E M A T I S Q A A
141 GCCTCCACGC AGGGCATCTC CGAAGACCTC TACAGCCGTT TAGTCGAAAT GGCCACTATC TCCCAAGCTG
    Sali
    .
    . Y A D L C N I P S T I I K G E K I Y N S Q T D .
211 CCTACGCCGA CCTGTGCAAC ATTCCGTCGA CTATTATCAA GGGAGAGAAA ATTTACAATT CTCAAACTGA
    BamHI
    .
    . I N G W I L R D D S S K E I I T V F R G T G S
281 CATTAACGGA TGGATCCTCC GCGACGACAG CAGCAAAGAA ATAATCACCG TCTTCCGTGG CACTGGTAGT
    D T N L Q L D T N Y T L T P F D T L P Q C N G C
351 GATACGAATC TACAACTCGA TACTAACTAC ACCCTCAGC CTTTCGACAC CCTACCACAA TGCAACGGTT
    . E V H G G Y Y I G W V S V Q D Q V E S L V K Q .
421 GTGAAGTACA CGGTGGATAT TATATTGGAT GGTCTCCGT CCAGGACCAA GTCGAGTCGC TTGTCAAACA
    . Q V S Q Y P D Y A L T V T G H X L G A S L A A
491 GCAGGTTAGC CAGTATCCGG ACTACGGCT GACCGTGACC GGCCACKCCC TCGGGCGCTC CCTGGCGGCA
    L T A A Q L S A T Y D N I R L Y T F G E P R S G
561 CTCACTGCCG CCCAGCTGTC TCGACATAC GACAACATCC GCCTGTACAC CTTTCGGCGAA CCGCGCAGCG
    XhoI
    .
    . N Q A F A S Y M N D A F Q A S S P D T T Q Y F .
631 GCAATCAGC CTTCCGCTCG TACATGAACG ATGCTTCCA AGCCTCGAGC CCAGATACGA CGCAGTATTT
    NcoI
    .
    . R V T H A N D G I P N L P P V E Q G Y A H G G

```

FIG._35B

```

701 CCGGGTCACT CATGCCAACG ACGGCATCCC AAACCTGCC CCGGTGGAGC AGGGGTACGC CCATGGCGGT
    V E Y W S V D P Y S A Q N T F V C T G D E V Q C
771 GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCAGAACCA CATTGTCTG CACTGGGGAT GAAGTGCAGT
    . C E A Q G G Q G V N N A H T T Y F G M T S G A .
841 GCTGTGAGGC CCAGGGCGGA CAGGTGTGA ATAATGCGCA CACGACTTAT TTTGGGATGA CGAGCGGAGC
    . C T W *
911 CTGTACATGG TGATCAGTCA TTTACGCCCTC CCCGAGTGTA CCAGGAAAGA TGGATGTCTT GGAGAGGGGG
981 CCGCGTAACC ACTGAAGGAT GAGCTGTAA GAAGCAGATC GTTCAAAACAT TTGGCAATAA AGTTTCTTAA
1051 GATTGAATCC TGTGCGCGGT CTTGCGATGA TTATCATATA ATTCTGTGTT AATTACGTTA AGCATGTAAT
1121 AATTAACATG TAATGCATGA CGTTATTTAT GAGATGGGTT TTTATGATTA GAGTCCCGCA ATTATACATT
1191 TAATACGCGA TAGAAAACAA AATATAGCGC GCAAAC TAGG ATAAATTATC GCGCGCGGTG TCATCTATGT
    HindIII
    ~~~~~
    ClaI          XbaI
    ~~~~~
1261 TACTAGATCG ATAAGCTTCT AGAGCGGCCG GTGGAGCTCC AATTCGCCCT ATAGTGAGTC GTATTACGCG
1331 CGTCACTGG CCGTCGTTT ACAACGTCTG GACTGGGAA ACCCTGGCGT TACCCAACTT AATCGCCTTG
1401 CAGCACATCC CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA GGGCCGCCCTT GATCGCCCTT CCCAACAGTT
1471 GCGCAGCCTG AATGGCGAAT GGGACGCGCC CTGTAGCGGC GCATTAAGCG CGCGGGGTGT GGTGGTTACG
1541 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCGC TTTCTTCCCT TCCTTCTCTG
1611 CCACGTTTCG CGGCTTTC CCGTCAAGCTC TAAATCGGGG GCTCCCTTTA GGGTTCGGAT TTAGTGCTTT
1681 ACGGCACCTC GACCCCAAAA AACTTGATTA GGGTGATGGT TCACGTAGTG GGCCATCGCC CTGATAGACG
1751 GTTTTTCGCC CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT GTTCCAAACT GGAACAACAC
1821 TCAACCCCTAT CTCGGTCTAT TCTTTTGATT TATAAGGGAT TTTGCCGATT TCGGCCTATT GGTAAAAAA
1891 TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAAACAA ATATTACGC TTACAAATTA GTGGGCACTT
1961 TTCGGGAAA TGTGCGCGGA ACCCTATTT GTTTATTTT CTAAATACAT TCAAAATATGT ATCCGCTCAT
2031 GAGACAATAA CCTGATAAA TGCTTCAATA ATATTGAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT
2101 GTCGCCCTTA TTCCCTTTT TCGGGCATTT TGCCTTCCTG TTTTTCCTCA CCCAGAAAACG CTGGTGAAG
2171 TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT
2241 CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTTCGTCT ATGTGGCGCG
2311 GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG
2381 TTGAGTACTC ACCAGTCACA GAAAGCATC TTACGGATGG CATGACAGTA AGAGAAATAT GCAGTGCTGC
2451 CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC
2521 GCTTTTTCG ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA
2591 TACCAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAC TATTAACCTGG
2661 CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA
```

FIG.-35C

```

2731 CTTCTGGCGT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC
2801 GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG
2871 TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCCT CACTGATTAA GCATTGGTAA
2941 CTGTCTAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACCTCA TTTTAAATTT AAAAGGATCT
3011 AGGTGAAGAT CCTTTTGTAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTTCGTTCC ACTGAGCGTC
3081 AGACCCCGTA GAAAAGATCA AAGATCTTC AAGATCTTC TTTTTCCTGC CGTAAATCTG CTGCTTGCAA
3151 ACAAAAAC CACCGCTACC AGCGGTGGTT TGTGTGCCG ATCAAGAGCT ACCAATCTTT TTTCCGAAGG
3221 TAACTGGCTT CAGCAGAGCG CAGATACCAA ATACTGTCTT TCTAGTGTAG CCGTAGTTAG GCCACCACTT
3291 CAAGAATCTT GTAGCACCGG GTTGGACTCA AGCAGTGG ATCCTGTATC CAGTGGCTGC TGCCAGTGGC
3361 GATAAGTCTG GTCTTACCGG GTTGGACTCA AGCAGTGG ATCCTGTATC TACCGGATAA GCGGCAGCGG TCGGGCTGAA
3431 CCGGGGGTTC GTGCACACAG CCCAGCTTGG AGCAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA
3501 GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGCG GACAGGTATC CCGTAAAGCG CAGGGTCGGA
3571 ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAACGCCCT GGTATCTTTA TAGTCCCTGC GGGTTTCGCC
3641 ACCTCTGACT TGAGCGTCCA TTTTGTGTAT GCTCGTCAAG GGGCGGAGC CTATGGAAA CTATGCCCTT
3711 CCGGGCTTTT TTACGGTTCC TGGCCTTTTG GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC
3781 GATTCGTGG ATAACCGTAT TACCGCTTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC
3851 GCAGCGAGTC AGTGAGCGAG GAAGCGGAG AGCGCCCAAT ACGCAAAACCG CCTCTCCCGG CGCGTTGGCC
3921 GATTCAATA TGCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGCA GTGAGCGCAA CGCAATTAAT
3991 GTGAGTTAGC TCACTCATTA GGCACCCACG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA
4061 TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCGCG CAATTAACCC

                                KpnI      XhoI
                                ~~~~~
4131 TCACTAAAGG GAACAAAAGC TGGGTACCGG GCCCCCCTC GAGGTCATTC ATATGCTTGA GAAAGAGATC
4201 GGGATAGTCC AAAATAAAAC AAAGGTAAGA TTACCTGGTC AAAAGTGAAA ACATCAGTTA AAAGGTGGTA
4271 TAAGTAAAT ATCGGTAATA AAAGGTGGCC CAAAGTGAAA TTTACTCTTT TCTACTATTA TAAAATTTGA
4341 GGATGTTTG TCGGTACTTT GATACGTCTAT TTTTGTATGA ATTGGTTTTT AAGTTTATTC GCGATTTGGA
4411 AATGCATATC TGTATTGAG TCGGTTTTTA AGTTCGTTGC TTTTGTAAAT ACAGAGGGAT TTGTATAAGA

                                EcoRI
                                ~~~~~
4481 AATATCTTTA AAAAACCCAT ATGCTAAATTT GACATAATTT TTGAGAAAAA TATATATTCA GGCGAATTCC
4551 ACAATGAACA ATAATAAGAT TAAAATAGCT TGCCCCCGTT GCAGCGATGG GTATTTTTTC TAGTAAAAATA
4621 AAAGATAAAC TTAGACTCAA AACATTTACA AAAACAACCC CTAAAGTCTT AAAGCCCAAA GTGCTATGCA
4691 CGATCCATAG CAAGCCCAGC CCAACCCAAC CCAACCCAGT CCACCCCACT GCAGCCCACT GGCAATAGT
4761 CTCACCCCC GGCACATATCA CCGTGAGTTG TCCGCACCCAC CGCACGCTCT GCAGCCCAAA AAAAAAAG
4831 AAAGAAAAAA AAGAAAAAGA AAAACAGCAG GTGGGTCCGG GTCGTGGGG CCGGAAAAAGC GAGGAGGATC

```

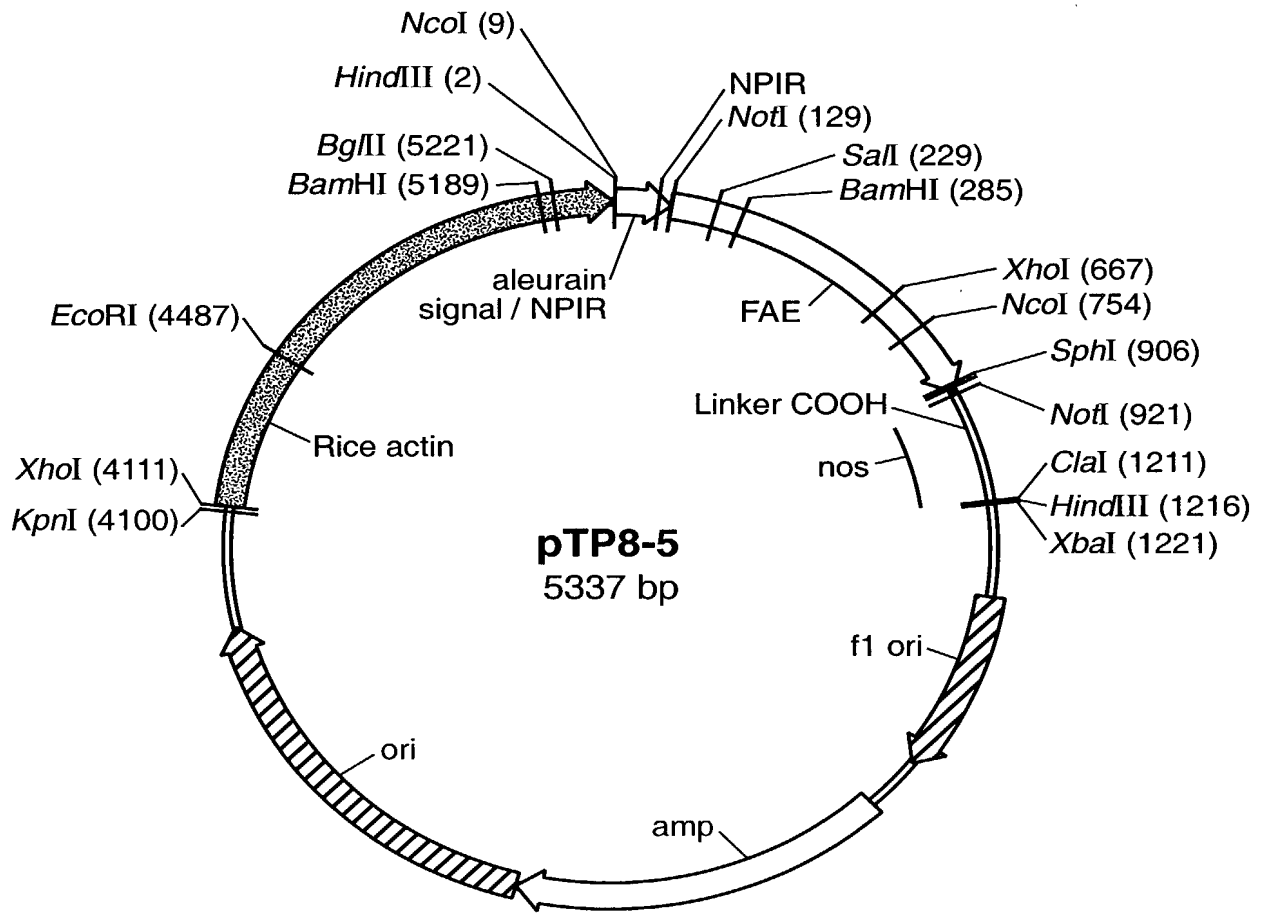
FIG.-35D

```
4901 GCGAGCAGCG ACGAGGCCCG GCCCTCCCTC CGCTTCCAAA GAAACGCCCC CCATCGCCAC TATATACATA
4971 CCCCCCCTC TCCTCCCATC CCCCACACC TACCACCAAC ACCACCAACA CTCTCTCCCC CCTCGCTGCC
5041 GGACGACGAG CTCCTCCCCC CTCCTCCCTCC GCCGCCGCCG GTAACCAACC CGCCCCCTCTC CTCCTTCTTTT
5111 CTCCTGTTT TTTTTCGTCT CCGTCTCGAT CTTTGGCCTT GGTAGTTTGG GTGGGCGAGA GCGGCTTCGT
5181 CGCCCAGATC GGTGCGCGGG AGGGCGGGA TCTCGCGGCT GCGTCTCTCCG GCGGTGAGTC GGCCCCGGATC
5251 CTCGCGGGGA ATGGGCTCT CGGATGTAGA TCTTCTTTCT TTCTTCTTTT TGTGGTAGAA TTTGAATCCC
5321 TCAGCATTTG TCATCGGTAG TTTTCTTTT CATGATTGT GACAAATGCA GCCTCGTGCG GAGCTTTTTT
5391 GTAGC
```

BamHI
~
BgIII
~~~~~

FIG..35E

55 / 154



**FIG.\_36A**

```

NcoI
~~~~~
HindIII
~~~~~
      M A H A R V L L L A L A V L A T A A V A V
1 AAGCTTACCA TGGCCACGC CCGGTCTCTC CTCCTGGCGC TCGCCGTGCT GCCACGGCC GCCGTGCGCG
      NotI
      . A S S S F A D S N P I R P V T D R A A A S T .
71 TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
      . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC
      SalI
      . . . . .
211 D L C N I P S T I I K G E K I Y N S Q T D I N G
      BamHI
      . W I L R D D S S K E I I T V F R G T G S D T N .
281 GATGGATCCT CCGGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACCTGGTA GTGATACGAA
      . L Q L D T N Y T L T P F D T L P Q C N G C E V
351 TCTACAACTC GATACTAAT ACACCTCAC GCCTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
      H G G Y I G W V S V Q D Q V E S L V K Q Q V S
421 CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAA CAGCAGGTTA
      . Q Y P D Y A L T V T G H X L G A S L A A L T A .
491 GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
      . A Q L S A T Y D N I R L Y T F G E P R S G N Q
561 CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGGCGAG CGGCAATCAG
      XhoI
      . . . . .
631 A F A S Y M N D A F Q A S S P D T T Q Y F R V T
      GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCAGATAC GACGCAGTAT TTCCGGGTCA
      NcoI
      . H A N D G I P N L P P V E Q G Y A H G G V E Y .
      ~~~~~
```

FIG.-36B



```

701 CTGATGCGAA CGACGGCATC CCAAACCTGC CCCCAGGTGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
 . W S V D P Y S A Q N T F V C T G D E V Q C C E
771 CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTTGTC TGCACTGGGG ATGAAGTGCA GTGCTGTGAG
 SphiI
    ~~~~~
    A Q G G Q G V N N A H T T Y F G M T S G A C T W
841 GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGC GCATGCACCT
    NotI
    ~~~~~
 . P V A A A *
911 GGCCGGTGGC GGCCGGGTAA CCACTGAAGG ATGAGCTGTA AAGAAGCAGA TCGTTCAAAC ATTGGCAAT
981 AAAGTTTCTT AAGATTGAAT CCTGTGCGG GTCCTGCGAT GATTATCATA TAAATTTCTGT TGAATTACGT
1051 TAAGCATGTA ATAATTAAACA TGTAAATGCAT GACGTTATTT ATGAGATGGG TTTTATATGAT TAGAGTCCCG
1121 CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG
 HindIII
    ~~~~~
    ClaI XbaI
    ~~~~~
1191 TGTATCTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC CTATAGTGAG
1261 TCGTATTACG CGCGCTCACT GGCGGTCTGT TTACAACGTC GTGACTGGGA AAACCTTGGC GTTACCCCAAC
1331 TTAATCGCCT TGCAGCACAT CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC
1401 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAG CGCGGCGGGT
1471 GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTCTTCC
1541 CTTCCCTTCT CGCCACGTTT TACGGCACC TCGACCCCAA AAAACTTGAT TAGGGTATG GTTCACGTA TGGGCCATCG
1611 ATTTAGTGCT TTACGGCACC CCGTTTTCG CCTTTTGAC TTGGAGTCCA CGTCTTTAA TAGTGGACTC TTGTTCCCAA
1681 CCTGATAGA CGGTTTTCG ACTCAACCTT ATCTCGGTCT ATTTCTTTGA TTTATAAGG ATTTTGCCGA TTTCGGCCTA
1751 CTGGAACAAC ACTCAACCTT ATCTCGGTCT ATTTCTTTGA TTTATAAGG ATTTTGCCGA TTTCGGCCTA
1821 TTGGTTAAA AATGAGCTGA TTTAACAACA AATTTAACA AAATATTAAC GCTTACAATT
1891 TAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAT ATTCAAATAT
1961 GTATCCGCTC ATGAGACAAT AACCTTGATA AATGCTTCAA TAATATTGA AAAGGAAGAG TATGAGTATT
2031 CAACATTTCC GTGTCGCCCT TATTCCTTTT TTTGCGGCAT TTTGCTTCC TGTTTTGTCT CACCCAGAAA
2101 CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA
2171 CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG
2241 CTATGTGGCG CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
2311 AGAATGACTT GGTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAAAT

```

FIG.\_36C

2381 ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAAACGAT CGGAGGACCG  
2451 AAGGAGCTAA CCGCTTTTTC GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC  
2521 TGAATGAAGC CATAACCAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA  
2591 ACTATTAACT GCGGAACCTAC TTAICTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GCGGATATAA  
2661 GTTGACAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTATTATGC TGATAAATCT GGAGCCCGTG  
2731 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC ATATGGATGA ACGAATATAG TGGTAAGCCC TCCCGTATCG TAGTTATCTA  
2801 CACGACGGGG AGTCAGGCAA CTATGGATGA CCAAGTTTAC TCATATATAC GAGATCGCTG AGATAGGTGC CTCACCTGATT  
2871 AAGCATTTGGT AACTGTCTAGA CCAAGTTTAC CCAAGTTTAC TCATATATAC TTTAGATTGA TTTTAAACCTT CATTTTAAAT  
2941 TTTAAAGGAT CTAGGTGAAG ATCCTTTTTG TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTCTT GCGCGTAATC  
3011 CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT CAAAGGATCT TCTTGAGATC CTTTGTGCTT GCGCGTAATC  
3081 TGCTGCTTGC AAACAAAATA ACCACCGCTA CCAGCGGTGG TTTGTGTTGGC GATCAACAGAG CTACCAACTC  
3151 TTTTTCGGA GGTAACTGGC TTCAGCAGAG CCGAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT  
3221 AGCCACCCAC TTCAAGAACT CTGTAGCACC GCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT  
3291 GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGTGTGACT CAAGACGATA GTTACCAGGAT AAGGCGCAGC  
3361 GGTGCGGCTG AACGGGGGT TCGTGACAC CAGCCAGCTT GGAGCGAAGG CGGACAGGTA TCCGGTAAGC  
3431 CCTACAGCGT GAGCTATGAG AAGCGCCAC GCTTCCCGAA GGGAGAAAGG CTGGTATCTT TATAGTCCCTG  
3501 GGCAGGGTCG GAACAGGAGA CCGCAGGAG GAGCTTCCAG GAGTTCGTA ATGCTGCTCA GGGGGCGGA GCTTATGGAA  
3571 TCGGGTTTCG CCACCTCTGA CTTGAGCGTC TTTTACGGTT CTTGGCCTT TTGCTCACAT GTTCTTTCTT  
3641 AACGCCAGC AACGCGGCTT TTTTACGGTT CTTGGCCTT TTGCTCACAT GTTCTTTCTT  
3711 GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCTT TTGAGTGAGC TGATACCGCT CGCCGCGAGC  
3781 GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCA ATACGCAAC CGCCTCTCCC  
3851 CGCGCGTTGG CCGATTCAAT AATGCAGCTG GCACGACAGG TTTCCCGACT GGAAGCGGG CAGTGAGCGC  
3921 AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA  
3991 TGTTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG AACACAGCTAT GACCATGATT ACGCCAAGCG

KpnI XhoI

4061 CGCAATTAACT CCTCACTAAA GGAACAAAA GCTGGGTACC GGGCCCCCCC TCAGAGGTCT TCATATGCTT  
4131 GAGAAGAGAG TCGGGATAGT CCAAAAATAA ACAAAGGTAA GATTACCTGG TCAAAAAGTGA AAACATCAGT  
4201 TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAGGTGG CCAAAAAGTGA AATTACTCT TTTCTACTAT  
4271 TATAAAAAAT GAGGATGTTT TGTCGGTACT TTGATACGTC ATTTTGTAT GAATTGGTTT TTAAGTTTAT  
4341 TCGCGATTGG GAAATGCATA TCTGTATTGG AGTCGGTTT TAAGTTCGTT GCTTTGTAA ATACAGAGGG  
4411 ATTTGTATAA GAAATATCTT TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA AATATATATT

EcoRI

4481 CAGGCGAATT CCACAATGAA CAATAATAAG ATTAAAAATAG CTTGCCCCCC TTGCAGCGAT GGGTATTTTT

FIG.\_36D

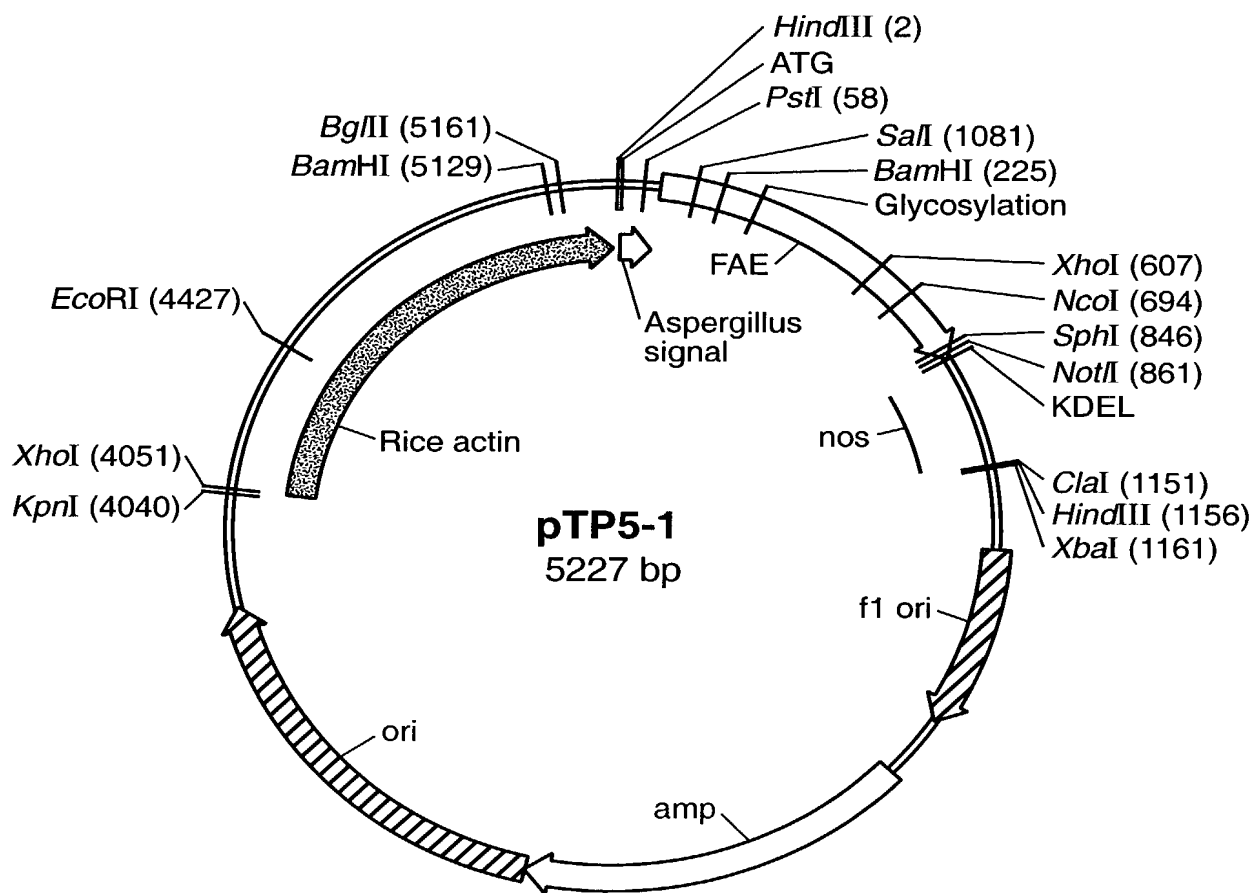
59 / 154

```
4551 TCTAGTAAA TAAAGATAA ACTTAGACTC AAAACATTTA CAAAAACAAC CCTAAAGTC CTAAAGCCCA
4621 AAGTGCTATG CACGATCCAT AGCAAGCCCA GCCAAACCCA ACCCAACCCA GTGCAGCCAA
4691 CTGGCAATA GTCTCCACC CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC TCGCAGCCAA
4761 AAAAAAATA AGAAGAAAAA AAAAGAAAAA AAAAGAAAAA AGGTGGGTCC GGTCTGTGGG GGCCGGAAAA
4831 GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGCCCTCCC TCCGCTTCCA AAAAAACGCC CCCATCGCC
4901 ACTATATACA TACCCCCCCC TCTCCTCCCA TCCCCCAAC CTACCACCA CCACCACCA CACCTCCTCC
4971 CCCCTCGCTG CCGGACGACG AGTCCTTCCC CCTCCCCCT CCGCCGCCGC CGGTAAACCAC CCCGCCCTC
5041 TCCTCTTTCT TTCTCCGTTT TTTTCTTCTG CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT GGTGGGGCGA
5111 GAGCGGCTTC GTCGCCAGA TCGGTGCGCG GGAGGGCGG GATCTCGCGG CTGGCGTCTC CGGGCGTGAG

 BamHI BglII
      ~~~~~          ~~~~~
5181 TCGGCCCGGA TCCTCGCGG GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG
5251 AATTGAATC CCTCAGCAT GTTCATCGGT AGTTTCTTCTT TTCATGATTT GTGACAAATG CAGCCTCGTG
5321 CGGAGCTTTT TTGTAGC
```

FIG.\_36E

60 / 154



**FIG.\_37A**

61 / 154

```

HindIII
~~~~~
 M K Q F S A K H V L A V V V T A G H A L A
1 AAGCTTAACA TGAAGCAGTT CTCGCCAAA CACGTCCTCG CAGTTGTGGT GACTGCAGGG CACGCCCTTAG
 . A S T Q G I S E D L Y S R L V E M A T I S Q A .
71 CAGCCTCTAC GCAAGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC
 Sali
      ~~~~~
   . A Y A D L C N I P S T I I K G E K I Y N S Q T
141 TGCCTACGCC GACCTGTGCA ACATTCCGTC GACTATTATC AAGGAGAGA AAATTACAA TTCTCAAAC T
      BamHI
      ~~~~~
D I N G W I L R D D S S K E I I T V F R G T G S
211 GACATTAACG GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACCTGGTA
 . D T N L Q L D T N Y T L T P F D T L P Q C N G .
281 GTGATACGAA TCTACAACTC GATACTAACT ACACCTTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG
 . C E V H G G Y I G W V S V Q D Q V E S L V K
351 TTGTGAAGTA CACGGTGGAT ATTATATTGG ATGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA
 Q Q V S Q Y P D Y A L T V T G H X L G A S L A A
421 CAGCAGGTTA GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG
 . L T A A Q L S A T Y D N I R L Y T F G E P R S .
491 CACTCACTGC CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG
 XhoI
      ~~~~~
   . G N Q A F A S Y M N D A F Q A S S P D T T Q Y
561 CGGCAATCAG GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT
      NcoI
      ~~~~~
F R V T H A N D G I P N L P P V E Q G Y A H G G
631 TTCCGGGTCA CTCATGCCAA CGACGGCATC CCAAACCTGC CCCGGGTGA GCAGGGGTAC GCCCATGGCG
 . V E Y W S V D P Y S A Q N T F V C T G D E V Q .
701 GTGTAGAGTA CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTTGTC TGCACCTGGG ATGAAGTGA
 . C C E A Q G G Q G V N N A H T T Y F G M T S G
771 GTGCTGTGAG GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGC

```

FIG..37B

```

SphI NotI
~~~~~                               ~~~~~
A  C  T  W  P  V  A  A  A  E  P  L  K  D  E  L  *
841 GCATGCACCT GGCCGGTCCG GCGCGCGGAA CCACTGAAGG ATGAGCTGTA AAGAAGCAGA TCGTTCAAAC
911 ATTTGGCAAT AAAGTTTCTT AAGATTGAAT CCTGTTGCCG GTCCTTCCGAT GATTATCATA TAAATTTCTGT
981 TGAATTACGT TAAGCATGTA ATAATTACAA TGTAATGCAT GACGTTATTT ATGAGATGGG TTTTATATGAT
1051 TAGAGTCCCG CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA
                                     HindIII
                                     ~~~~~
 ClaI XbaI
                                     ~~~~~
1121 TCGCGCGCGG TGTCACTCTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTTCGCC
1191 CTATAGTGAG TCGTATTACG CGCGCTCACT GGCGTCTGTT TTACAACGTC GTGACTGGGA AAACCCCTGGC
1261 GTTACCCAAAC TTAATCGCCT TGCAGCACAT CCCCCTTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA
1331 CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAG
1401 CGCGCGGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC
1471 GCTTCTTCC CTTCTTCTTCT CGCCACGTTT TACGGCACC TCGACCCCAA AAAACTTGAT TAGGTGATG GTTCAACGTAG
1541 TAGGTTCCG ATTTAGTGCT TTACGGCACC CCGTTTTTCG CCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC
1611 TGGGCCATCG CCTGATAGA CGGTTTTTCG ACTCAACCTT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG ATTTTGCCGA
1681 TTGTTCCAA CTGGAACAA AATGAGCTGA TTTTACAAA AATTAACGCG AATTTTAAAC AAATAATTAAC
1751 TTTTCGGCCTA TTGGTTAAA TAGGTGGCAC TTTTCGGGA AATGTGCGG GAACCCCTAT TTGTTTATTT TTCTAAATAC
1821 GCTTACAAAT TAGGTGGCAT GTATCCGCTC ATGAGACAAT AACCTTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG
1891 ATTCAAATAT GTATCCGCTC GTGTCCGCTT GTGTCCGCTT TATTCCTTTT TTTGCGGCAT TTTGCTTCC ACAGTGGGT
1961 TATGAGTATT CAACATTTCC CGCTGGTGAA AGTAAAGAT GCTGAAGATC AGTTGGGTGC ACAGTGGGT TACATCGAAC
2031 CACCCAGAAA CGCTGGTGAA CAGCGGTAG ATCCTTGAGA GTTTTCGCC CCGTATTTGAC CGCGGCAAG AGCAACTCGG TCGCCGCATA
2101 TGGATCTCAA TAAAGTTCTG CTATGTGGCG CGGTATTATC GGTATTGAG TCACCAAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG
2171 CACTATTCTC AGAATGACTT GGTGAGTAC GGCATAACCA TGAGTGATAA CACTGCGGCC GGGGATCATG TAACTCGCCT TGACACACGAT
2241 TAAGAGAATT ATGCAGTGCT GCCATAACCA CCGCTTTTTT GCACAAACATG GACGAGCGTG ACACACGAT GCCTGTAGCA ATGGCAACAA
2311 CGGAGGACCG AAGGAGCTAA CCGCTTTTTT CATAACAAAC GACGAGCGTG TTTCTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA
2381 GAACCGGAGC TGAATGAAGC CATAACAAAC GACGAGCGTG TTTCTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA
2451 CGTTGCGCAA ACTATTAACT GCGGAACCTAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT
2521 GCGCGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT
2591 GGAGCCGGTG AGCGTGGGTC TCGCGGTATC ATTCAGCAC ATTCAGCAC TGGGGGCCAG TGGTAAAGCCC TCCCCTATCG
2661
```

FIG.\_37C

63 / 154

```
2731 TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC
2801 CTCACCTGATT AAGCATTTGGT AACTGTTCAGA CCAAGTTTAC TCATATATATAC TTTAGATTGA TTTAAAACCTT
2871 CATTTTAAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTGG ATAATCTCAT GACCAAAATC CTTAAACGTG
2941 AGTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT
3011 GCGCGTAATC TGCTGCTTGC AAACAAAAAA GGTAACCTGGC TTCAGCAGAG CCACCGGTGG TTTGTTTGCC GGATCAAGAG
3081 CTACCAACTC TTTTTCGGAA AGGCCACCAC TTCAAGAACT GTGTCTTACC GCTTACATAC AAATACTGTC CTTCTAGTGT
3151 AGCCGTAGTT AGCCGAGTG GCTGCCAGTG AACGGGGGT TCGTGCACAC AGCCAGCTT GGAGCGAAGC ACCTACACCG
3221 ACCAGTGGCT GCTGCCAGTG GCTGCCAGTG AACGGGGGT TCGTGCACAC AGCCAGCTT GGAGCGAAGC ACCTACACCG
3291 AAGGCGCAGC GGTGCGGCTG CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA
3361 AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA
3431 TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT
3501 TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTGTG ATGCTCGTCA GGGGGCGCGA
3571 GCCTATGGAA AAACGCCAGC AACGCGGCTT TTTTACGGTT CTTGGCCCTT TGCTGCACAT
3641 GTTCTTTTCTT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCTT TTGAGTGAGC TGATACCGCT
3711 CGCCGCAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAAC
3781 CGCCTCTCCC CGCGCGTTGG CCGATTCAAT AATGCAGCTG GCACGACAGG TTTCCCGACT GGAAGCGGGG
3851 CAGTGAGCGC AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT
3921 CCGGCTCGTA TGTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG AACAGCTAT GACCATGATT

                                     KpnI
                                     ~~~~~~
3991 ACGCCAAGCG CGCAATTAAAC CCTCACTAAA GGGAAACAAA GCTGGGTACC GGGCCCCCCC TCGAGGTGAT
4061 TCATATGCTT GAGAAGAGAG TCGGGATAGT CCAAAATAAA ACAAAAGTAA GATTACCTGG TCAAAAAGTGA
4131 AAACATCAGT TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAAGGTGG CCCAAAGTGA AATTACTCT
4201 TTTCTACTAT TATAAAAATT GAGGATGTTT TGTCGGTACT TTGATACGTC ATTTTGTGAT GAAATTGGTTT
4271 TTAAGTTTAT TCGCGATTG GAAATGCATA TCTGTATTG AGTCGGTTTT TAAGTTCGTT GCTTTTGTA
4341 ATACAGAGGG ATTTGTATTA GAAATATCTT TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA

 EcoRI
                                     ~~~~~~
4411 AATATATATT CAGGCGAATT CCACAATGAA CAATAATAAG ATTAAAATAG CTTGCCCCCG TTGCAGCGAT
4481 GGGTATTTTT TCTAGTAAAA TAAAAGATAA ACTTAGACTC AAAACATTTA CAAAACAAC CCCTAAAGTC
4551 CTAAGCCCCA AAGTGCTATG CACGATCCAT AGCAAGCCCA GCCCAACCCA ACCCAACCCA
4621 GTGCAGCCCA CTGGCAATA GTCTCCACC CCGGCACTAT CACCGTGAGT TGTCCGCAAC ACCGACGTC
4691 TCGCAGCCCA AAAAAAAA AGAAGAAA AAAAGAAA GAAAACAGC AGGTGGGTCC GGTGCTGTGG
```

FIG.-37D

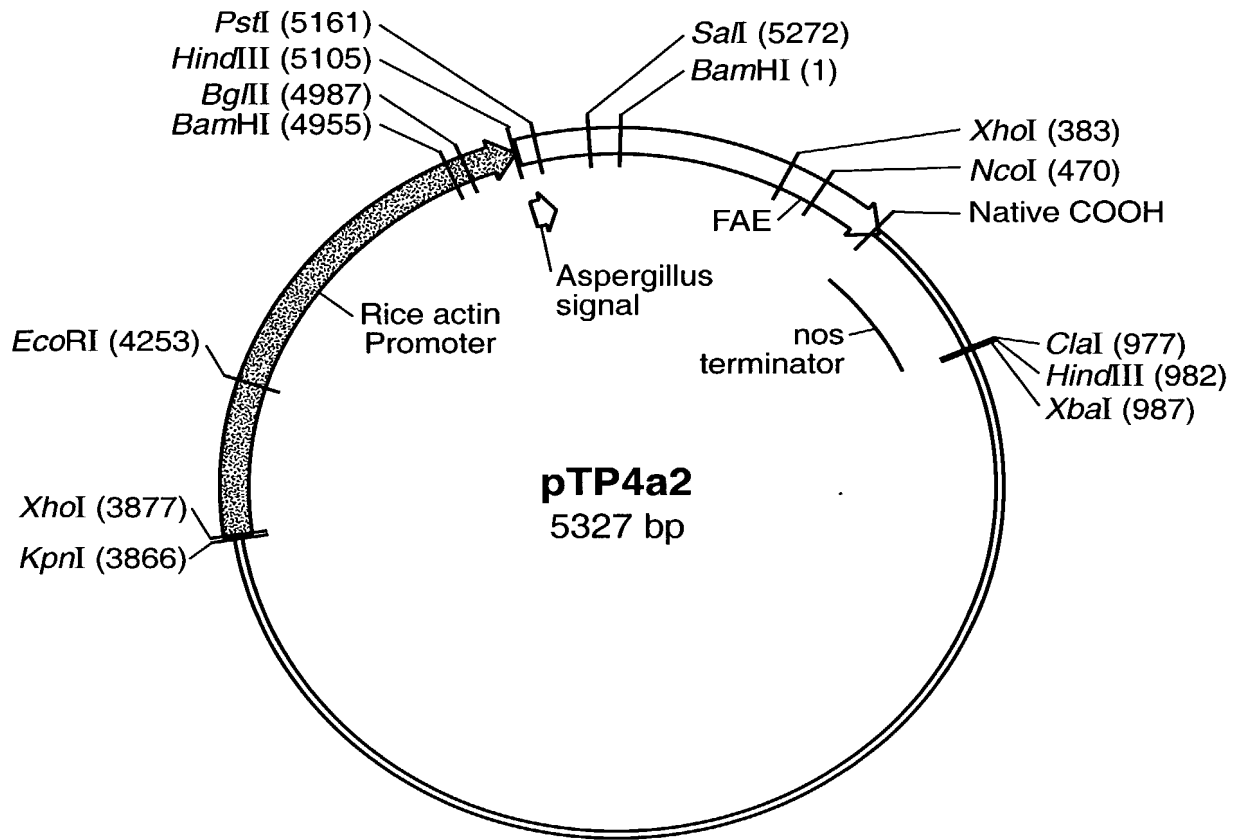
64 / 154

```
4761 GGC CGG AAA GCG AGG AGG A TCG CGA GCG TCG AGG AGG CGG AGG AGG CGG AGG AGG CGG AGG AGG CGG
4831 CCC CAT CGC ACT ATA TACA TAC CCCCCC TCT CCT FCCA TCCCCC AAC CCT ACC ACCA CCACC ACCAC
4901 CAC CTC CTC CCC CTC GCTG CCG GAC GAC AGC TCC TCCC CCT TCCCCCT CCG CCG CCG CCG CCG TAA ACCAC
4971 CCC GCCCC TC TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT TCT
5041 GGT GGG CGA GAG GGT TC GCG CCGA GCG CCGA GCG CCGA GCG CCGA GCG CCGA GCG CCGA GCG CCGA GCG
      BamHI
      ~~~~~
5111 CGG CGT GAG TCG GCC CGG TCT CGC GCG TCT CGG GGT GAAT GGG GGT CTC GGA TGA GAT CTC TCT TCT
5181 TTT GTG GTAG AAT TGA ATC CCT CAG CAT GTT CAT CCGT GTT CAT CCGT GTT CAT CCGT GTT CAT CCGT
5251 CAG CCGT CCGT CCG AGC TTT TTGTAG
 BglII
      ~~~~~
      TCG GGT TCG GGT GAT CTC GCG GGT GAT CTC GCG GGT GAT CTC GCG GGT GAT CTC GCG GGT TCT
```

FIG. 37E



65 / 154



**FIG.\_38A**

66 / 154

BamHI

~~~~~  
 . I L R D D S S K E I I T V F R G T G S D T N L
 1 GATCCTCCGC GACGACAGCA GCAAAGAAAT AATCACCCTC TTCCGTGGCA CTGGTAGTGA TACGAATCTA
 Q L D T N Y T L T P F D T L P Q C N G C E V H G
 71 CAACTCGATA CTAACATACAC CCTCAGCCTT TTCGACACCC TACCACAATG CAACGGTTGT GAAGTACACG
 . G Y Y I G W V S V Q D Q V E S L V K Q Q V S Q .
 141 GTGGATATTA TATTGGATGG GTCCTCCGTCC AGGACCAAGT CGAGTCGCTT GTCAAACAGC AGGTTAGCCA
 . Y P D Y A L T V T G H X L G A S L A A L T A A
 211 GTATCCGGAC TACGCGCTGA CCGTGACCGG CCACKCCCTC GCGCCTCC TGGCGGCACT CACTGCCGCC
 Q L S A T Y D N I R L Y T F G E P R S G N Q A F
 281 CAGCTGTCTG CGACATACGA CAACATCCGC CTGTACACCT TCGGCGAACC GCGCAGCGGC AATCAGGCCT
 XhoI

~~~~~  
 . A S Y M N D A F Q A S S P D T T Q Y F R V T H .  
 351 TCGCGTCGTA CATGAACGAT GCCTTCCAAG CCTCGAGCCC AGATACGACG CAGTATTTC GGGTCACTCA  
 NcoI

~~~~~  
 . A N D G I P N L P P V E Q G Y A H G G V E Y W
 421 TGCCAACGAC GGCATCCCAA ACCTGCCCCC GGTGAGCAG GGTACGCC ATGGCGGTGT AGAGTACTGG
 S V D P Y S A Q N T F V C T G D E V Q C C E A Q
 491 AGCGTTGATC CTTACAGCGC CCAGAACACA TTTGTCTGCA CTGGGGATGA AGTGCAGTGC TGTGAGGCC
 . G G Q G V N N A H T T Y F G M T S G A C T W * .
 561 AGGCGGACA GGGTGTGAAT AATGCCACA CGACTTATT TGGGATGACG AGCGGAGCCT GTACATGGTG
 . *
 631 ATCAGTCATT TCAGCCTCCC CGAGTGTAAC AGGAAAGATG GATGTCTTGG AGAGGGGCC GCGTAACCCAC
 701 TGAAGGATGA GCTGTAAAGA AGCAGATCGT TCAAACATTT GGCAATAAAG TTTCTTAAGA TTGAATCCTG
 771 TTGCCGGTCT TCGGATGATT ATCATATAAT TTCTGTTGAA TTACGTTAAG CATGTAATAA TTAACATGTA
 841 ATGCATGACG TTATTATGA GATGGGTTTT TATGATTAGA GTCCCGCAAT TATACATTA ATACGGGATA
 ClaI
 911 GAAACAATAA TATAGCGGC AACTAGGAT AAATTATCGC GCGCGGTGTC ATCTATGTTA CTAGATCGAT
 XbaI

~~~~~  
 HindIII  
 ~~~~~

FIG. 38B

67 / 154

981 AAGCTTCTAG AGCGGCCGGT GGAGCTCCAA TTCGCCCTAT AGTGAGTCGT ATTACGCGCG CTCACATGGCC
1051 GTCGTTTTAC AACGTCGTGA CTGGGAAAAC CCTGGCGTTA CCCAACTTAA TCGCCTTGCA GCACATCCCC
1121 CTTTCGCCAG CTGGCGTAAT AGCGGCGCTT GTAGCGGCGC ATTAAGCGCG CCGCCCTTCC CAACAGTTGC GCAGCCTGAA
1191 TGGCGAATGG GACGCGCCCT AGCGGCCCTT AGCGGCCCTT CTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG
1261 GCTACACTTG CTAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA
1331 GCTTTCCCGG CTTGATTAGG GTGATGGTTC ACGTAGTGG GACTCTTGT TCCAAACTGG AACAACTC AACCCATATCT
1401 CCCCCAAAAC CTTGATTAGG GTGATGGTTC ACGTAGTGG GACTCTTGT TCCAAACTGG AACAACTC AACCCATATCT
1471 TTGACGTTGG AGTCCACGTT CTTTAATAGT GACTCTTGT TCCAAACTGG AACAACTC AACCCATATCT
1541 CGGTCATATC TTTTGATTTA TAAGGGATTT TGCCGATTTG GGCCTATTGG TTAATAAAATG AGCTGATTTA
1611 AAAAAAATTT AACGCGAATT TTAACAAAAT ATTAACGCTT ATTAACGCTT ACAATTTAGG TGGCCTTTT CGGGGAAAATG
1681 TGCGCGGAAC CCTATTTGT TTAATTTTCT AAATACATTC AAATATGTAT CCGCTCATGA GACAATAACC
1751 CTGATAAATG CTTCAATAAT ATTGAAAAAG GAAGAGTATG AGTATCAAC ATTTCCGTGT GGTGAAAGTA AAAGATGCTG
1821 CCGCTTTTGG CGGCATTTTG CCGCATTTTG CTTCTCTGTT TTTGCTCACC CAGAAACGCT GGTGAAAGTA AAAGATGCTG
1891 AAGATCAGTT GGTGACACGA GTGGGTACCA CAATGATGAG CACTTTTAA GTTCTGCTAT GTGGCGCGGT ATTATCCCGT
1961 TCGCCCCGAA GAACGTTTTC CAATGATGAG CACTTTTAA GTTCTGCTAT GTGGCGCGGT ATTATCCCGT
2031 ATTGACGCGG GCAAGAGCA ACTCGGTCGC CGCATACACT ATTCTCAGAA TGACTTGGTT GAGTACTCAC
2101 CAGTCACAGA AAAGCATCTT ACGGATGGCA TGACAGTAAG AGAATTATGC AGTGCTGCCA TAACCATGAG
2171 TGATAACACT GCGGCCAAT TACTTCTGAC AACGATCGGA GGACCGAAGG AGCTAACCGC TTTTITGACAC
2241 AACATGGGGG ATCATGTAAC TCGCCTTGAT CGTTGGGAA CCGAGCTGAA TGAAGCCATA CCAACGACG
2311 AGCGTGACAC CACGATGCCT GTAGCAATGG CAACAACGTT GCGCAAACTA TTAACCTGGC AACTACTTAC
2381 TCTAGCTTCC CGGCAACAAT TAATAGACTG GATGGAGCG GATAAAGTTG CAGGACCACT TCTGCGCTCG
2451 GCCCTTCCGG CTGGCTGGTT TATTGCTGAT AAATCTGGAG CCGGTGAGCG TGGGTCTCGC GGTATCATTG
2521 CAGCACTGGG GCCAGATGGT AAGCCCTCCC GTATCGTAGT TATCTACACG ACGGGAGTC AGGCAACTAT
2591 GGATGAACGA AATAGACAGA TCGCTGAGAT AGGTGCCCTCA CTGATTAAAGC ATTGGTAACT GTCAGACCAA
2661 GTTTACTCAT ATATACTTTA GATTGATTTA AAACCTCAT TTTAATTTAA AAGGATCTAG GTGAAGATCC
2731 TTTTGTGATA TCTCATGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC TGAGCGTCAG ACCCGTAGA
2801 AAAGATCAA GGTCTTCTT GAGATCCCTT TTTTCTGCGC GATACTGCT TCCGAAAGTA ACTGGCTTCA
2871 CCGTACCAG CCGTGGTTTG TTTGCCGGAT CAAGAGCTAC CAACTCTTTT TCCGAAAGTA ACTGGCTTCA
2941 GCAGAGCGCA GATACCAAAT ACTGTCTTTC TAGTGAGCC GTAGTTAGGC CACCACTTCA AGAATCTGT
3011 AGCACCGCCT ACATACCTCG CTCTGCTAAT CCGGATAAG CGCAGCGGT GGGCTGAAACG GGGGTTCGT
3081 CTTACCGGGT TGGACTCAAG ACGATAGTTA CCGGATAAG CGCAGCGGT GGGCTGAAACG GGGGTTCGT
3151 GCACACAGCC CAGCTTGGAG CGAACGACCT ACACCGAAT GAGATACCTA CAGCGTAGC TATGAGAAAG
3221 CGCCACGCTT CCGAAGGGA GAAAGCGGA CAGGTATCCG GTAAGCGGCA GGTTCGGAAC AGGAGAGCGC
3291 ACGAGGGAGC TTCCAGGGGG AAACGCTTGG TATCTTTATA GTCTGTCTGG GTTTCGCCAC CTCTGACTTG
3361 AGCGTCGATT TTTGTGATGC TCGTCAGGGG GCGGAGCCT ATGGAATAAC GCCAGCAACG CGGCCCTTTT

FIG. 38C

```
3431 ACGGTTCTG GCCTTTTGCT GGCTTTTTC TCACATGTTT TTTCTGCGT TATCCCTCGA TTCTGTGGAT
3501 AACCGTATTA CCGCCTTTGA GTGAGCTGAT ACCGCTCGC GCAGCCGAAC GACCGAGCGC AGCGAGTCAG
3571 TGAGCGAGGA AGCGGAAGAG CGCCCAATAC GCAACCCGCC TCTCCCGCGG CGTTGGCCGA TTCATTAATG
3641 CAGCTGGCAC GACAGGTTTC CCGACTGGAA AGCGGCGAGT GAGCGCAACG CAATTAATGT GAGTTAGCTC
3711 ACTCATTAGG CACCCAGGC TTTACACTTT ATGCTTCGG CTCGTATGTT GTGTGGAATT GTGAGCGGAT
3781 AACAAATTCA CACAGGAAAC AGCTATGACC ATGATTACGC CAAGCGCGCA ATTAACCCCTC ACTAAAGGGA

      KpnI
      ~~~~~
3851 ACAAAGCTG GGTACCGGGC CCCCCTCGA GGTCATTCA ATGCTTGAGA AGAGAGTCGG GATAGTCCAA
3921 AATAAAACAA AGGTAAGATT ACCTGGTCAA AAGTGAAAC ATCAGTTAAA AGGTGGTATA AGTAAATAT
3991 CGGTAATAAA AGGTGGCCCA AAGTGAAATT TACTCTTTTC TACTATTATA AAAATTGAGG ATGTTTGTG
4061 GGTACTTTGA TACGTCATTT TGTGATGAAT TGGTTTTTAA GTTTATTCGC GATTTGGAAA TGCATATCTG
4131 TATTTGAGTC GGTTTTAAAG TTCGTTGCTT TTGTAATAC AGAGGGATT TTATAAGAAA TATCTTTAAA

      EcoRI
      ~~~~~
4201 AAACCCATAT GCTAATTITGA CATAATTTT GAGAAAAATA TATATTCAGG CGAATTCCAC AATGAACAAT
4271 AATAAGATTA AATAGCTTG CCCCCTTGC AGCGATGGGT ATTTTCTTA GTAAAAATA AGATAAACTT
4341 AGACTCAAAA CATTTACAAA AACAAACCTT AAAGTCTTAA AGCCCAAAGT GCTATGCACG ATCCATAGCA
4411 AGCCAGGCC AACCCAAACC AACCCAAACC ACCCCAGTGC AGCCAACTGG CAAATAGTCT CCACCCCGG
4481 CACTATCACG GTGAGTTGTC CGCACCAACG CACGCTCGC AGCCAAAAA AAAAAAGAA AGAAAAAAA
4551 GAAAAAGAAA AACAGCAGGT GGTCCGGGT CGTGGGGGCC GGAAGAGCGA GGAGGATCGC GAGCAGCGAC
4621 GAGGCCCGGC CCTCCCTCCG CTTCCAAAGA AACGCCCCCC ATCGCCACTA TATACATACC CCCCCCTCTC
4691 CTCCCATCCC CCAACCCCTA CCACCAACAC CACCACCAAC TCCTCCCGCC TCGCTGCCGG ACGACGAGCT
4761 CCTCCCCCCT CCCCCTCCG CGCCGCCGGT AACCAACCCG CCCCTCTCCT CTTCTTTTCT CCGTTTCTT
4831 TTTCGTCCTG GTCTCGATCT TTGGCCCTTG TAGTTTGGGT GGGCGAGAGC GGCCTCGTCG CCCAGATCGG

      BamHI
      ~~~~~
4901 TGCGCGGGAG GGGCGGGATC TCGCGGCTGG CGTCTCCGG CGTGAGTCGG CCCGGATCCT CGCGGGGAAT

      BglII
      ~~~~~
4971 GGGGCTCTCG GATGTAGATC TTCTTTCTTT CTCTTTTTC TGGTAGAATT TGAATCCCTC AGCATTTGTC

      HindIII
      ~~~~~
5041 ATCGGTAGTT TTTCTTTTCA TGATTGTGA CAAATGCAGC CTCGTGCGGA GCTTTTTTGT AGCAAGCTTA
```

FIG. 38D

69 / 154

```

PstI
~~~~~
M K Q F S A K H V L A V V T A G H A L A A S .
5111 ACATGAAGCA GTTCTCCGCC AACACGTCC TCGCAGTTGT GGTGACTGCA GGGCAGCCCT TAGCAGCCTC
. T Q G I S E D L Y S R L V E M A T I S Q A A Y
5181 TACGCAAGGC ATCTCCGAAG ACCTCTACAG CCGTTTAGTC GAAATGGCCA CTATCTCCCA AGCTGCCCTAC
Sali
~~~~~
A D L C N I P S T I I K G E K I Y N S Q T D I N
5251 GCCGACCTGT GCAACATTCC GTCGACTATT ATCAAGGGAG AGAAAATTTA CAATTCTCAA ACTGACATTA
B
. G W
5321 ACGGATG
```

FIG.-38E

70 / 154

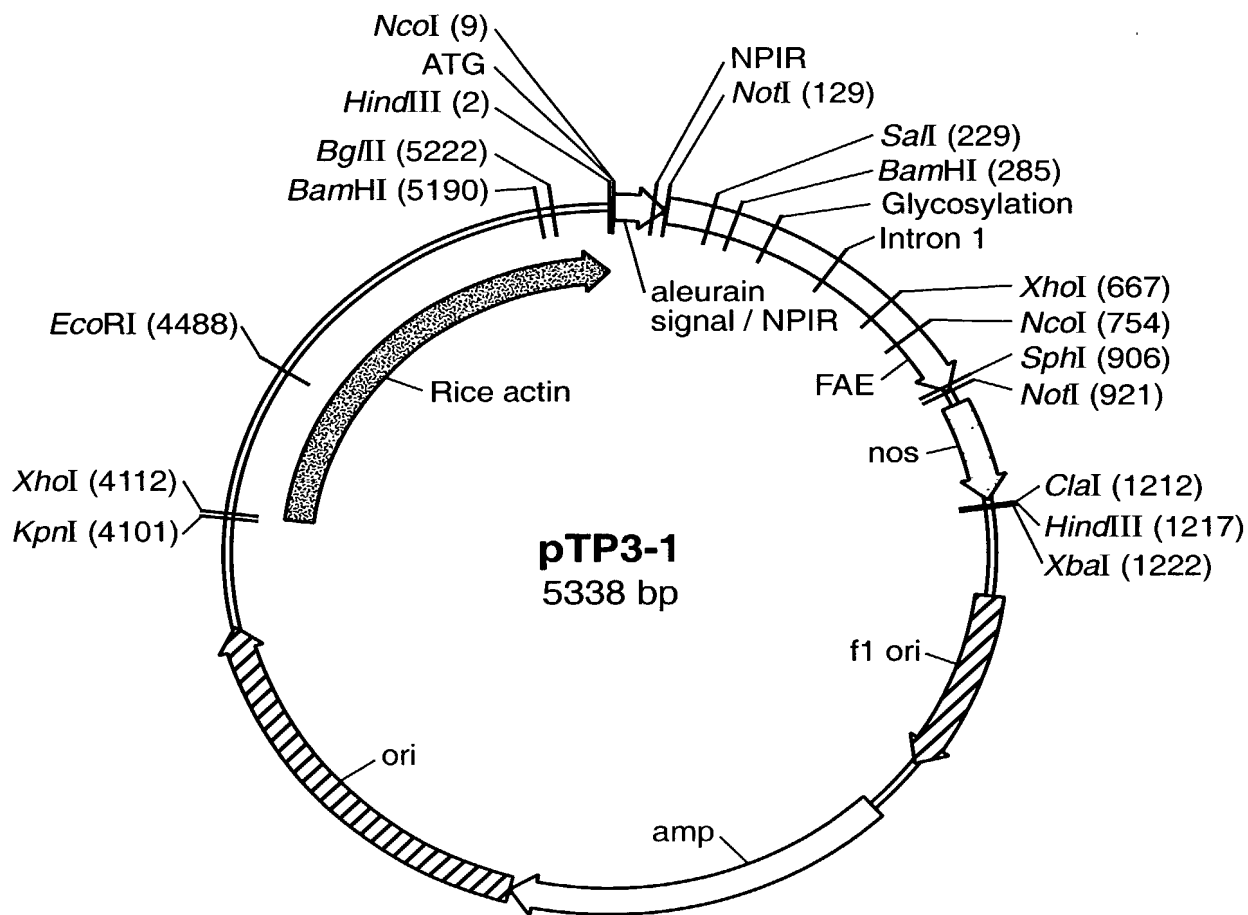


FIG._39A

71 / 154

```

NcoI
~~~~~
HindIII
~~~~~
      M A H A R V L L L A L A V L A T A A V A V
1 AAGCTTACCA TGGCCACGC CCGGTCTCTC CTCCTGGCGC TCGCCGTGCT GCCACGGCC GCCGTGCGCG

      . A S S S F A D S N P I R P V T D R A A A S T .
71 TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCGATCCG GCCGTCACC GACCGGCGG CCGCCTCCAC
  . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC

Sali
~~~~~
      D L C N I P S T I I K G E K I Y N S Q T D I N G
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGAGAGA AAATTACAA TTCTCAAAC T GACATTAACG

BamHI
~~~~~
      . W I L R D D S S K E I I T V F R G T G S D T N .
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCAC TGGTA GTGATACGAA

Glycosylation
~~~~~
      . L Q L D T N Y T L T P F D T L P Q C N G C E V
351 TCTACAAC TC GATACTAACT ACACCTCAC GCCTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
      H G G Y Y I G W V S V Q D Q V E S L V K Q Q V S
421 CACGGTGGAT ATTATATTGG ATGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAA CAGCAGGTTA
      . Q Y P D Y A L T V T G H X L G A S L A A L T A .
491 GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
      . A Q L S A T Y D N I R L Y T F G E P R S G N Q
561 CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
```

FIG._39B

72 / 154

```

XhoI
~~~~~
631  A F A S Y M N D A F Q A S S P D T T Q Y F R V T
    GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA

NcoI
~~~~~
701  . H A N D G I P N L P P V E Q G Y A H G G V E Y .
    CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
    . W S V D P Y S A Q N T F V C T G D E V Q C C E
771  CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTGTC TGCACCTGGG ATGAAGTGCA GTGCTGTGAG

SphI
~~~~~
841  A Q G G Q G V N N A H T T Y F G M T S G A C T W
    GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGC GCATGCACCT

NotI
~~~~~
KDEL
~~~~~
911  . P V A A A E T T E G *
    GCGCGTCCG GCGCGCGGAA ACCACTGAAG GATGAGCTGT AAAGAAAGCAG ATCGTTCAA CATTGGCAA
981  TAAAGTTCT TAAGATTGAA TCCTGTGCG GGTCTTGGCA TGATTATCAT ATAATTCTG TTGAATTACG
1051 TTAAGCATGT AATAATTAA ATGTAATGCA TGACGTTATT TATGAGATGG GTTTTATGA TTAGAGTCCC
1121 GCAATTATAC ATTTAATACG CGATAGAAA CAAAATATAG CGCGCAAAT AGGATAAATT ATCGCGCGCG

HindIII
~~~~~
ClaI
~~~~~
XbaI
~~~~~
1191 GTGTCATCTA TGTACTAGA TCGATAAGCT TCTAGAGCGG CCGGTGGAGC TCCAATTGCG CCTATAGTGA
1261 GTCGTATTAC GCGCGCTCAC TGGCCGTCGT TTACAAACGT CGTGACTGGG AAAACCCCTGG CGTTACCCAA
1331 CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC
1401 CTTCCCAACA GTTGGCGAGC CTGAATGGCG AATGGGACGC GCCCTGTAGC GGCGCATTA GCGCGGCGGG

```

FIG.-39C

73 / 154

1471 TGTGGTGGTT ACGGCAGCG TGACCGCTAC ACTTGCCAGC GCCTAGCGC CCGCTCCTTT CGCTTCTTTC
 1541 CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCGGTCAAG CTCATAATCG GGGCTCCCTT TTAGGGTTCC
 1611 GATTTAGTGC TTTACGGCAC TTTACGGCAC AAAAAGTTGA TTAGGGTGAT GGTTCACGTA GTGGGCCATC
 1681 GCCCTGATAG ACGGTTTTTC CACCTTTGAC GCCCTTTGAC GTTGAGTCC ACGTTCTTTA ATAGTGACT CTTGTTCCAA
 1751 ACTGGAACAA CACTCAACCC CACTCAACCC TATCTCGGTC TATCTTTTG AATTATAAGG GATTTGCGG ATTTGCGCCT
 1821 AATTGGTTAA AATGAGCTG AATTAACAA AATTAAACG GAATTTTAA CAAATATAA AAAATATAA CGCTTACAAT
 1891 TTAGGTGGCA CTTTTCGGG AATGTGCGC AAATGTGCGC GGAACCCCTA TTTGTTTAT TTTCTAAATA CATTCAAATA
 1961 TGTATCCGCT CATGAGACAA TAACCTTGAT TAACCTTGAT AAATGTGCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
 2031 TCAACATTTT CGTGTGCGC TATTTCCCTT TATTTCCCTT TTTTGCGGCA TTTTGCCCTC CTGTTTTGTC TCACCCAGAA
 2101 ACGCTGGTGA AAGTAAAGA TGCTGAAGAT TGCTGAAGAT CAGTTGGTG CACGAGTGG TTACATCGAA CTGGATCTCA
 2171 ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG TTTTCCAATG ATGAGCACTT TTAAGTTCT
 2241 GCTATGTGGC GCGGTATTAT CCGGTATTAT CTCACCATG CTCACCATG ATCTTACGGA TGGCATGACA GTAAAGAAAT
 2311 CAGAAATGACT TGGTTGAGTA TGGTTGAGTA CTCACCATG CTCACCATG ATCTTACGGA TGGCATGACA GTAAAGAAAT
 2381 TATGAGTGC TGCCATAACC TGCCATAACC ATGAGTGATA ACCTGCGC ACAATTAATA GACTGGATGG AGCGGATAA
 2451 GAAGGAGCTA ACCGCTTTT TGCACAAACAT GGGGATCAT GTAACCTGCC CTGATCGTTG GGAACCCGAG
 2521 CTGAATGAAG CCATACCAA CGACGAGCGT CTTTCCCGCA CTTTCCCGCA TGCCTGTAGC AATGGCAACA ACGTTGCGCA
 2591 AACTATTAACT TGGCGAACTA CTTACTCTAG CTTTCCCGCA ACAATTAATA GACTGGATGG AGCGGATAA
 2661 AGTTGCAGGA CCATTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT
 2731 GAGCGTGGGT CTCGCGGTAT CATTCAGCA CTGGGCGCAG ATGGTAAGCC CTCCTGTATC GTAGTTATCT
 2801 ACACGACGGG GAGTCAGGA ACTATGGATG AACGAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
 2871 TAAGCATGG TAACTGTGAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAAACT TCATTTTAA
 2941 TTTAAAAGGA TCTAGGTGAA GATCCTTTT GATAATCTCA TGACCAAAAT CCTTAAACGT GAGTTTTCGT
 3011 TCCACTGAGC GTCAGACCC GTAGAAAAGA TCAAAGGATC ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GTACCAACT
 3081 CTGCTGCTTG CAAACAAAAG AACCACCGCT CTTACAGACA GCGCAGATAC CAAATACTGT CTTCTAGTG TAGCCGTAGT
 3151 CTTTTCCTGA AGGTAACCTG CTTCAAGAAC TCTGTAGCAC CGCTTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
 3221 TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCTTACATA CCTCGCTCTG TCAAGACGAT AGTTACCGA TAAAGCGCAG
 3291 TGCTGCCAGT GGCATAAAGT TCGTGTGACA CAGCCACGCT TGGAGCGAAC GACCTACACC GAACTGAGAT
 3361 CGGTCGGGCT GAACGGGGG TCGTGTGACA CAGCCACGCT TGGAGCGAAC GACCTACACC GAACTGAGAT
 3431 ACCTACAGCG TGAGCTATGA GAAAGCGCA CGCTTCCGA AGGAGAAAG GCGGACAGGT ATCCGGTAAG
 3501 CGGCAGGCT GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGAAACG CTTGGTATCT TTATAGTCTT
 3571 GTCGGGTTTC GCCACCTCTG ACTTGAGCGT CGATTTTGT GATGCTCGTC AGGGGGCGG AGCCTATGGA
 3641 AAAACGCCAG CAACGCGGCC TTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTCTTTCC
 3711 TGCGTTATCC CCTGATCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC TCGCCGCGAGC
 3781 CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC AATACGCAAA CCGCTCTCTC
 3851 CCGCGCGTTG GCCGATTCTA TAATGCAGCT GGCACGACAG GTTTCCCGAC TGGAAAAGCGG GCAGTGAGCG

FIG. 39D

74 / 154

```
3921 CAACGCAATT AATGTGAGTT AGTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT
3991 ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG GAAACAGCTA TGACCATGAT TACGCCAAGC
      ~~~~~
      KpnI
4061 GCGCAATTAA CCTCACTAA AGGGAACAAA AGTGGGTAC CGGGCCCCC CTCGAGGTCA TTCATATGCT
4131 TGAGAAGAGA GTCGGGATAG TCCAAAATAA AACAAAGTA AGATTACCTG GTCAAAAGTG AAAACATCAG
4201 TTAAAAGGTG GTATAAGTAA AATATCGGTA ATAAAAGTG GCCAAAGTG AAATTACTC TTTTCTACTA
4271 TTATAAAAAA TGAGGATGTT TTGTCCGTAC TTTGATACGT CATTTTGTG TGAATTGGTT TTTAAAGTTA
4341 TTCGCCGATT GGAATGCAAT ATCTGTATTT GAGTCGGTTT TTAAGTTCGT TGCCTTTGTA AATACAGAGG
4411 GATTGTGATA AGAAATATCT TTAATAAAACC CATATGCTAA TTTGACATAA TTTTTGAGAA AAATATATAT
      ~~~~~
      ECORI
4481 TCAGGCGAAT TCCACAATGA ACAATAATAA GATTAAATA GCTTGCCCC GTTGCAGCGA TGGGTATTTT
4551 TTCTAGTAAA ATAAAAGATA AACTTAGACT CAAAACATTT ACAAAAACAA CCCCTAAAGT CCTAAAAGCCC
4621 AAAGTGCTAT GCACGATCCA TAGCAAGCCC AGCCCAACCC AACCCACCCC AACCCACCCC AGTGCAGCCA
4691 ACTGGCAAAAT AGTCTCCACC CCCGGCACTA TCACCGTGAG TTGTCCGCAC CACCGCACGT CTCGCAGCCA
4761 AAAAAAATAA AAGAAAGAAA AAAAAAGAAA AGAAAACAG CAGGTGGGTC CGGGTCGTGG GGGCCGGAAA
4831 AGCGAGGAGG ATCGCGAGCA GCGACGAGGC CCGGCCCTCC CTCCGCTTCC AAAGAAACGC CCCCATCGC
4901 CACTATATAC ATACCCCCC CTCTCCTCCC ATCCCCCCTCC CCTACCACC ACCACCACCA CCACCTCCTC
4971 CCCCCTCGCT GCCGGACGAC GAGCTCCTCC CCCCTCCCC TCCGCCGCCG CCGGTAACCA CCCC GCCCT
5041 CTCCTCTTTC TTTCTCCGTT TTTTCTTTCG TCTCGGTCCT GATCTTTGGC CTGGGTAGTT TGGGTGGGCG
5111 AGAGCGGCTT CGTCGCCCAG ATCGGTGCGC GGGAGGGGCG GGATCTCGCG GCTGGCGTCT CCGGGCGTGA
      ~~~~~
      BamHI
5181 GTCGGCCCCG ATCCTCGCGG GGAATGGGCG TCTCGGATGT AGATCTTCTT TCTTCTTCTT TTTTGTGGTA
5251 GAATTGGAAT CCTCAGCAT TGTTCATCGG TAGTTTCTT TTTTCATGAT TGTGACAAAT GCAGCCTCGT
5321 GCGGAGCTTT TTTGTAGC
```

FIG. 39E

75 / 154

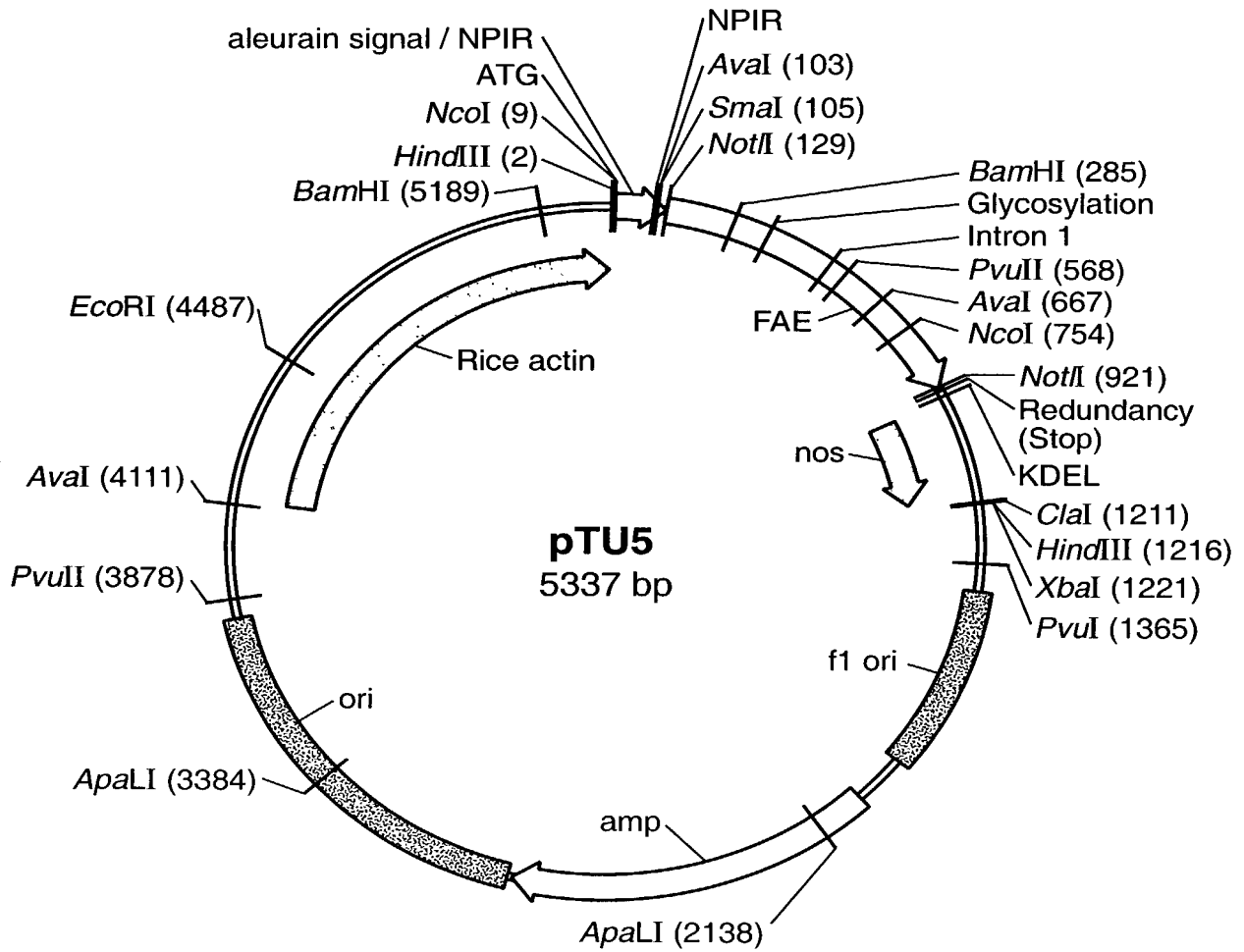


FIG. 40A

76 / 154

HindIII NcoI
 ~~~~~

1 AAGCTTACCA TGGCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT  
 TTCGAATGGT ACCGGGTGCG GGCAGCAGGAG GAGGACCGCG AGCGGCACGA

51 GGCCACGGCC GCCGTCGCCG TCGCCTCCTC CTCCTCCTTC GCCGACTCCA  
 CCGGTGCCCG CGGCAGCGGC AGCGGAGGAG GAGGAGGAAG CGGCTGAGGT

**SmaI**  
 ~~~~~

AvaI
 ~~~~~

**NotI**  
 ~~~~~

101 ACCCGGGCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC GCAGGGCATC
 TGGGCCCCGGC CGGGCAGTGG CTGGCGCGCC GGCAGGAGGTG CGTCCCCTAG

151 TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC
 AGGCTTCTGG AGATGTCGGC AAATCAGCTT TACCGGTGAT AGAGGGTTTCG

201 TGCCTACGCC GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA
 ACGGATGCGG CTGGACACGT TGTAAGGCAG CTGATAATAG TTCCCTCTCT

BamHI
 ~~~~~

251 AAATTTACAA TTCTCAAAC TACATTAACG GATGGATCCT CCGCGACGAC  
 TTTAAATGTT AAGAGTTTGA CTGTAATTGC CTACCTAGGA GGCCTGCTG

301 AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCAGTGGTA GTGATACGAA  
 TCGTCGTTTC TTTATTAGTG GCAGAAGGCA CCGTGACCAT CACTATGCTT

351 TCTACAAC TC GATACTAACT ACACCTCAC GCCTTTCGAC ACCCTACCAC  
 AGATGTTGAG CTATGATTGA TGTGGGAGTG CGGAAAGCTG TGGGATGGTG

401 AATGCAACGG TTGTGAAGTA CACGGTGGAT ATTATATTGG ATGGGTCTCC  
 TTACGTTGCC AACACTTCAT GTGCCACCTA TAATATAACC TACCCAGAGG

451 GTCCAGGACC AAGTCGAGTC GCTTGTCAA CAGCAGGTTA GCCAGTATCC  
 CAGGTCCTGG TTCAGCTCAG CGAACAGTTT GTCGTCCAAT CGGTCATAGG

501 GGA CTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG  
 CCTGATGCGC GACTGGCACT GGCCGGTGMG GGAGCCGCGG AGGGACCGCC

**PvuII**  
 ~~~~~

551 CACTCACTGC CGCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC
 GTGAGTGACG GCGGGTCGAC AGACGCTGTA TGCTGTTGTA GGCGGACATG

601 ACCTTCGGCG AACCGCGCAG CGGCAATCAG GCCTTCGCGT CGTACATGAA
 TGGAAGCCGC TTGGCGCGTC GCCGTTAGTC CGGAAGCGCA GCATGTACTT

AvaI
 ~~~~~

651 CGATGCCTTC CAAGCCTCGA GCCAGATAC GACGCAGTAT TTCCGGGTCA  
 GCTACGGAAG GTTCGGAGCT CGGGTCTATG CTGCGTCATA AAGGCCAGT

**FIG.\_40B**

77 / 154

|      |              |             |            |            |            |
|------|--------------|-------------|------------|------------|------------|
| 701  | CTCATGCCAA   | CGACGGCATC  | CCAAACCTGC | CCCCGGTGGA | GCAGGGGTAC |
|      | GAGTACGGTT   | GCTGCCGTAG  | GGTTTGGACG | GGGGCCACCT | CGTCCCCATG |
|      | NcoI         |             |            |            |            |
|      | ~~~~~        |             |            |            |            |
| 751  | GCCCATGGCG   | GTGTAGAGTA  | CTGGAGCGTT | GATCCTTACA | GCGCCCAGAA |
|      | CGGGTACCGC   | CACATCTCAT  | GACCTCGCAA | CTAGGAATGT | CGCGGGTCTT |
| 801  | CACATTTGTC   | TGCACTGGGG  | ATGAAGTGCA | GTGCTGTGAG | GCCCAGGGCG |
|      | GTGTAAACAG   | ACGTGACCCC  | TACTTCACGT | CACGACACTC | CGGGTCCCGC |
| 851  | GACAGGGTGT   | GAATAATGCG  | CACACGACTT | ATTTTGGGAT | GACGAGCGGC |
|      | CTGTCCACAC   | CTTATTACGC  | GTGTGCTGAA | TAAAACCCTA | CTGCTCGCCG |
|      | NotI         |             |            |            |            |
|      | ~~~~~        |             |            |            |            |
| 901  | GCATGCACCT   | GGCCGGTCGC  | GGCCGCGGAA | CCACTGAAGG | ATGAGCTGTA |
|      | CGTACGTGGA   | CCGGCCAGCG  | CCGGCGCCTT | GGTGACTTCC | TACTCGACAT |
| 951  | AAGAAGCAGA   | TCGTTCAAAC  | ATTTGGCAAT | AAAGTTTCTT | AAGATTGAAT |
|      | TTCTTCGTCT   | AGCAAGTTTG  | TAAACCGTTA | TTTCAAAGAA | TTCTAACTTA |
| 1001 | CCTGTTGCCG   | GTCTTGCGAT  | GATTATCATA | TAATTTCTGT | TGAATTACGT |
|      | GGACAACGGC   | CAGAACGCTA  | CTAATAGTAT | ATTAAAGACA | ACTTAATGCA |
| 1051 | TAAGCATGTA   | ATAATTAACA  | TGTAATGCAT | GACGTTATTT | ATGAGATGGG |
|      | ATTCGTACAT   | TATTAATTGT  | ACATTACGTA | CTGCAATAAA | TACTCTACCC |
| 1101 | TTTTTATGAT   | TAGAGTCCCG  | CAATTATACA | TTTAATACGC | GATAGAAAAC |
|      | AAAAATACTA   | ATCTCAGGGC  | GTTAATATGT | AAATTATGCG | CTATCTTTTG |
| 1151 | AAAATATAGC   | GCGCAAACCTA | GGATAAATTA | TCGCGCGCGG | TGTCATCTAT |
|      | TTTTTATATCG  | CGCGTTTGAT  | CCTATTTAAT | AGCGCGCGCC | ACAGTAGATA |
|      | XbaI         |             |            |            |            |
|      | ~~~~~        |             |            |            |            |
|      | ClaI HindIII |             |            |            |            |
|      | ~~~~~        |             |            |            |            |
| 1201 | GTTACTAGAT   | CGATAAGCTT  | CTAGAGCGGC | CGGTGGAGCT | CCAATTCGCC |
|      | CAATGATCTA   | GCTATTTCGAA | GATCTCGCCG | GCCACCTCGA | GGTTAAGCGG |
| 1251 | CTATAGTGAG   | TCGTATTACG  | CGCGCTCACT | GGCCGTCGTT | TTACAACGTC |
|      | GATATCACTC   | AGCATAATGC  | GCGCGAGTGA | CCGGCAGCAA | AATGTTGCAG |
| 1301 | GTGACTGGGA   | AAACCCTGGC  | GTTACCCAAC | TTAATCGCCT | TGCAGCACAT |
|      | CACTGACCCT   | TTTGGGACCG  | CAATGGGTTG | AATTAGCGGA | ACGTCGTGTA |
|      | PvuII        |             |            |            |            |
|      | ~~~~~        |             |            |            |            |
| 1351 | CCCCCTTTTCG  | CCAGCTGGCG  | TAATAGCGAA | GAGGCCCGCA | CCGATCGCCC |
|      | GGGGGAAAGC   | GGTCGACCGC  | ATTATCGCTT | CTCCGGGCGT | GGCTAGCGGG |
| 1401 | TTCCCAACAG   | TTGCGCAGCC  | TGAATGGCGA | ATGGGACGCG | CCCTGTAGCG |
|      | AAGGGTTGTC   | AACGCGTCGG  | ACTTACCGCT | TACCCTGCGC | GGGACATCGC |

FIG.\_40C

78 / 154

|                |                           |                          |                          |                           |                           |
|----------------|---------------------------|--------------------------|--------------------------|---------------------------|---------------------------|
| 1451           | GCGCATTAAAG<br>CGCGTAATTC | CGCGGCGGGT<br>GCGCCGCCCA | GTGGTGGTTA<br>CACCACCAAT | CGCGCAGCGT<br>GCGCGTCGCA  | GACCGCTACA<br>CTGGCGATGT  |
| 1501           | CTTGCCAGCG<br>GAACGGTTCG  | CCCTAGCGCC<br>GGGATCGCGG | CGCTCCTTTC<br>GCGAGGAAAG | GCTTTCTTCC<br>CGAAAGAAGG  | CTTCCTTTCT<br>GAAGGAAAGA  |
| 1551           | CGCCACGTTT<br>GCGGTGCAAG  | GCCGGCTTTC<br>CGGCCGAAAG | CCCGTCAAGC<br>GGGCAGTTTC | TCTAAATCGG<br>AGATTTAGCC  | GGGCTCCCTT<br>CCCGAGGGAA  |
| 1601           | TAGGGTTCCG<br>ATCCCAAGGC  | ATTTAGTGCT<br>TAAATCACGA | TTACGGCACC<br>AATGCCGTGG | TCGACCCCAA<br>AGCTGGGGTT  | AAAACCTTGAT<br>TTTTGAACTA |
| 1651           | TAGGGTGATG<br>ATCCCACTAC  | GTTACAGTAG<br>CAAGTGCATC | TGGGCCATCG<br>ACCCGGTAGC | CCCTGATAGA<br>GGGACTATCT  | CGGTTTTTTCG<br>GCCAAAAAGC |
| 1701           | CCCTTTGACG<br>GGGAAACTGC  | TTGGAGTCCA<br>AACCTCAGGT | CGTTCTTTAA<br>GCAAGAAATT | TAGTGGACTC<br>ATCACCTGAG  | TTGTTCCAAA<br>AACAAGGTTT  |
| 1751           | CTGGAACAAC<br>GACCTTGTTG  | ACTCAACCCT<br>TGAGTTGGGA | ATCTCGGTCT<br>TAGAGCCAGA | ATTCCTTTTGA<br>TAAGAAAAC  | TTTATAAGGG<br>AAATATTCCC  |
| 1801           | ATTTTGCCGA<br>TAAAACGGCT  | TTTCGGCCTA<br>AAAGCCGGAT | TTGGTTAAAA<br>AACCAATTTT | AATGAGCTGA<br>TTACTCGACT  | TTTAACAAAA<br>AAATTGTTTT  |
| 1851           | ATTTAACGCG<br>TAAATTGCGC  | AATTTTAACA<br>TTAAAATTGT | AAATATTAAC<br>TTTATAATTG | GCTTACAATT<br>CGAATGTAA   | TAGGTGGCAC<br>ATCCACCGTG  |
| 1901           | TTTTCGGGGA<br>AAAAGCCCC   | AATGTGCGCG<br>TTACACGCGC | GAACCCCTAT<br>CTTGGGGATA | TTGTTTATTT<br>AACAAATAAA  | TTCTAAATAC<br>AAGATTTATG  |
| 1951           | ATTCAAATAT<br>TAAGTTTATA  | GTATCCGCTC<br>CATAGGCGAG | ATGAGACAAT<br>TACTCTGTTA | AACCCTGATA<br>TTGGGACTAT  | AATGCTTCAA<br>TTACGAAGTT  |
| 2001           | TAAATATTGAA<br>ATTATAACTT | AAAGGAAGAG<br>TTTCCTTCTC | TATGAGTATT<br>ATACTCATAA | CAACATTTCC<br>GTTGTAAAGG  | GTGTCGCCCT<br>CACAGCGGGA  |
| 2051           | TATTCCTTTT<br>ATAAGGGAAA  | TTTGCGGCAT<br>AAACGCCGTA | TTTGCCTTCC<br>AAACGGAAGG | TGTTTTTTGCT<br>ACAAAAACGA | CACCCAGAAA<br>GTGGGTCTTT  |
| ApaLI<br>~~~~~ |                           |                          |                          |                           |                           |
| 2101           | CGCTGGTGAA<br>GCGACCACTT  | AGTAAAAGAT<br>TCATTTTCTA | GCTGAAGATC<br>CGACTTCTAG | AGTTGGGTGC<br>TCAACCCACG  | ACGAGTGGGT<br>TGCTCACCCA  |
| 2151           | TACATCGAAC<br>ATGTAGCTTG  | TGGATCTCAA<br>ACCTAGAGTT | CAGCGGTAAG<br>GTCGCCATTC | ATCCTTGAGA<br>TAGGAACTCT  | GTTTTTCGCC<br>CAAAAGCGGG  |
| 2201           | CGAAGAACGT<br>GCTTCTTGCA  | TTTCCAATGA<br>AAAGGTACT  | TGAGCACTTT<br>ACTCGTGAAA | TAAAGTTCTG<br>ATTTCAGAC   | CTATGTGGCG<br>GATACACCGC  |
| 2251           | CGGTATTATC<br>GCCATAATAG  | CCGTATTGAC<br>GGCATAACTG | GCCGGGCAAG<br>CGGCCCGTTC | AGCAACTCGG<br>TCGTTGAGCC  | TCGCCGCATA<br>AGCGGCGTAT  |
| 2301           | CACTATTCTC<br>GTGATAAGAG  | AGAATGACTT<br>TCTTACTGAA | GGTTGAGTAC<br>CCAACTCATG | TCACCAGTCA<br>AGTGGTCAGT  | CAGAAAAGCA<br>GTCTTTTCGT  |

FIG.\_40D

79 / 154

|      |                           |                          |                          |                           |                          |
|------|---------------------------|--------------------------|--------------------------|---------------------------|--------------------------|
| 2351 | TCTTACGGAT<br>AGAATGCCTA  | GGCATGACAG<br>CCGTACTGTC | TAAGAGAATT<br>ATTCTCTTAA | ATGCAGTGCT<br>TACGTCACGA  | GCCATAACCA<br>CGGTATTGGT |
| 2401 | TGAGTGATAA<br>ACTCACTATT  | CACTGCGGCC<br>GTGACGCCGG | AACTTACTTC<br>TTGAATGAAG | TGACAACGAT<br>ACTGTTGCTA  | CGGAGGACCG<br>GCCTCCTGGC |
| 2451 | AAGGAGCTAA<br>TTCTCTCGATT | CCGCTTTTTT<br>GGCGAAAAAA | GCACAACATG<br>CGTGTTGTAC | GGGGATCATG<br>CCCCTAGTAC  | TAACTCGCCT<br>ATTGAGCGGA |
| 2501 | TGATCGTTGG<br>ACTAGCAACC  | GAACCGGAGC<br>CTTGGCCTCG | TGAATGAAGC<br>ACTTACTTCG | CATACCAAAC<br>GTATGGTTTG  | GACGAGCGTG<br>CTGCTCGCAC |
| 2551 | ACACCACGAT<br>TGTGGTGCTA  | GCCTGTAGCA<br>CGGACATCGT | ATGGCAACAA<br>TACCGTTGTT | CGTTGCGCAA<br>GCAACGCGTT  | ACTATTAACT<br>TGATAATTGA |
| 2601 | GGCGAACTAC<br>CCGCTTGATG  | TTACTCTAGC<br>AATGAGATCG | TTCCCGGCAA<br>AAGGGCCGTT | CAATTAATAG<br>GTTAATTATC  | ACTGGATGGA<br>TGACCTACCT |
| 2651 | GGCGGATAAA<br>CCGCCTATTT  | GTTGCAGGAC<br>CAACGTCCTG | CACTTCTGCG<br>GTGAAGACGC | CTCGGCCCTT<br>GAGCCGGGAA  | CCGGCTGGCT<br>GGCCGACCGA |
| 2701 | GGTTTATTGC<br>CCAAATAACG  | TGATAAATCT<br>ACTATTTAGA | GGAGCCGGTG<br>CCTCGGCCAC | AGCGTGGGTC<br>TCGCACCCAG  | TCGCGGTATC<br>AGCGCCATAG |
| 2751 | ATTGCAGCAC<br>TAACGTCGTG  | TGGGGCCAGA<br>ACCCCGGTCT | TGGTAAGCCC<br>ACCATTCGGG | TCCCGTATCG<br>AGGGCATAGC  | TAGTTATCTA<br>ATCAATAGAT |
| 2801 | CACGACGGGG<br>GTGCTGCCCC  | AGTCAGGCAA<br>TCAGTCCGTT | CTATGGATGA<br>GATACCTACT | ACGAAATAGA<br>TGCTTTTATCT | CAGATCGCTG<br>GTCTAGCGAC |
| 2851 | AGATAGGTGC<br>TCTATCCACG  | CTCACTGATT<br>GAGTGACTAA | AAGCATTGGT<br>TTCGTAACCA | AACTGTCAGA<br>TTGACAGTCT  | CCAAGTTTAC<br>GGTTCAAATG |
| 2901 | TCATATATAC<br>AGTATATATG  | TTTAGATTGA<br>AAATCTAACT | TTTAAAACTT<br>AAATTTTGAA | CATTTTAAAT<br>GTAAAAATTA  | TTAAAAGGAT<br>AATTTTCCTA |
| 2951 | CTAGGTGAAG<br>GATCCACTTC  | ATCCTTTTTT<br>TAGGAAAAAC | ATAATCTCAT<br>TATTAGAGTA | GACCAAAATC<br>CTGGTTTTAG  | CCTTAACGTG<br>GGAATTGCAC |
| 3001 | AGTTTTTCGTT<br>TCAAAAGCAA | CCACTGAGCG<br>GGTGACTCGC | TCAGACCCCG<br>AGTCTGGGGC | TAGAAAAGAT<br>ATCTTTTCTA  | CAAAGGATCT<br>GTTTCCTAGA |
| 3051 | TCTTGAGATC<br>AGAACTCTAG  | CTTTTTTTCT<br>GAAAAAAGA  | GCGCGTAATC<br>CGCGCATTAG | TGCTGCTTGC<br>ACGACGAACG  | AAACAAAAAA<br>TTTGTTTTTT |
| 3101 | ACCACCGCTA<br>TGGTGGCGAT  | CCAGCGGTGG<br>GGTCGCCACC | TTTGTTTGCC<br>AAACAAACGG | GGATCAAGAG<br>CCTAGTTCTC  | CTACCAACTC<br>GATGGTTGAG |
| 3151 | TTTTTCCGAA<br>AAAAAGGCTT  | GGTAACTGGC<br>CCATTGACCG | TTCAGCAGAG<br>AAGTCGTCTC | CGCAGATACC<br>GCGTCTATGG  | AAATACTGTC<br>TTTATGACAG |
| 3201 | CTTCTAGTGT<br>GAAGATCACA  | AGCCGTAGTT<br>TCGGCATCAA | AGGCCACCAC<br>TCCGGTGGTG | TTCAAGAACT<br>AAGTTCTTGA  | CTGTAGCACC<br>GACATCGTGG |
| 3251 | GCCTACATAC<br>CGGATGTATG  | CTCGCTCTGC<br>GAGCGAGACG | TAATCCTGTT<br>ATTAGGACAA | ACCAGTGGCT<br>TGGTCACCGA  | GCTGCCAGTG<br>CGACGGTCAC |

FIG.\_40E

80 / 154

3301 GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT  
 CGCTATTTCAG CACAGAATGG CCAACCTGA GTTCTGCTAT CAATGGCCTA

ApaLI

~~~~~

3351 AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT
 TTCCGCGTCG CCAGCCCGAC TTGCCCCCA AGCACGTGTG TCGGGTCGAA

3401 GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG
 CCTCGCTTGC TGGATGTGGC TTGACTCTAT GGATGTCGCA CTCGATACTC

3451 AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC
 TTTCGCGGTG CGAAGGGCTT CCCTCTTTCC GCCTGTCCAT AGGCCATTCG

3501 GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC
 CCGTCCCGAC CTTGTCTCTT CCGGTGCTCC CTCGAAGGTC CCCCTTTGCG

3551 CTGGTATCTT TATAGTCCTG TCGGGTTTTCG CCACCTCTGA CTTGAGCGTC
 GACCATAGAA ATATCAGGAC AGCCCAAAGC GGTGGAGACT GAACTCGCAG

3601 GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC
 CTA AAAACAC TACGAGCAGT CCCCCGCTT CGGATACCTT TTTGCGGTCTG

3651 AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT
 TTGCGCCGGA AAAATGCCAA GGACCGGAAA ACGACCGGAA AACGAGTGTA

3701 GTTCTTTTCTT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT
 CAAGAAAGGA CGCAATAGGG GACTAAGACA CCTATTGGCA TAATGGCGGA

3751 TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG
 AACTCACTCG ACTATGGCGA GCGGCGTCGG CTTGCTGGCT CGCGTCGCTC

3801 TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC CGCCTCTCCC
 AGTCACTCGC TCCTTCGCCT TCTCGCGGGT TATGCGTTTG GCGGAGAGGG

PvuII

~~~~~

3851 CGCGCGTTGG CCGATTCAAT AATGCAGCTG GCACGACAGG TTTCCCGACT  
 GCGCGCAACC GGCTAAGTAA TTACGTCGAC CGTGCTGTCC AAAGGGCTGA

3901 GGAAAGCGGG CAGTGAGCGC AACGCAATTA ATGTGAGTTA GCTCACTCAT  
 CCTTTCGCCC GTCACTCGCG TTGCGTTAAT TACACTCAAT CGAGTGAGTA

3951 TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA TGTGTGTGG  
 ATCCGTGGGG TCCGAAATGT GAAATACGAA GGCCGAGCAT ACAACACACC

4001 AATTGTGAGC GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT  
 TTAACACTCG CCTATTGTTA AAGTGTGTCC TTTGTCGATA CTGGTACTAA

4051 ACGCCAAGCG CGCAATTAAC CCTCACTAAA GGAACAAAA GCTGGGTACC  
 TGCGGTTTCG GCGTTAATTG GGAGTGATTT CCCTTGTTTT CGACCCATGG

AvaI

~~~~~

4101 GGGCCCCCCC TCGAGGTCAT TCATATGCTT GAGAAGAGAG TCGGGATAGT
 CCCGGGGGGG AGCTCCAGTA AGTATACGAA CTCTTCTCTC AGCCCTATCA

FIG._40F

81 / 154

4151	CCAAAATAAA	ACAAAGGTAA	GATTACCTGG	TCAAAAGTGA	AAACATCAGT
	GGTTTTATTT	TGTTTTCCATT	CTAATGGACC	AGTTTTTCACT	TTTGTAGTCA
4201	TAAAAGGTGG	TATAAGTAAA	ATATCGGTAA	TAAAAGGTGG	CCCAAAGTGA
	ATTTTCCACC	ATATTCATTT	TATAGCCATT	ATTTTCCACC	GGGTTTCACT
4251	AATTTACTCT	TTTCTACTAT	TATAAAAATT	GAGGATGTTT	TGTCGGTACT
	TTAAATGAGA	AAAGATGATA	ATATTTTTTAA	CTCCTACAAA	ACAGCCATGA
4301	TTGATACGTC	ATTTTTGTAT	GAATTGGTTT	TTAAGTTTAT	TCGCGATTTG
	AACTATGCAG	TAAAAACATA	CTTAACCAAA	AATTCAAATA	AGCGCTAAAC
4351	GAAATGCATA	TCTGTATTTG	AGTCGGTTTT	TAAGTTCGTT	GCTTTTGTA
	CTTTACGTAT	AGACATAAAC	TCAGCCAAAA	ATTCAAGCAA	CGAAAAACATT
4401	ATACAGAGGG	ATTTGTATAA	GAAATATCTT	TAAAAAACCC	ATATGCTAAT
	TATGTCTCCC	TAAACATATT	CTTTATAGAA	ATTTTTTTGGG	TATACGATTA
				EcoRI	
				~~~~~	
4451	TTGACATAAT	TTTTGAGAAA	AATATATATT	CAGGCGAATT	CCACAATGAA
	AACTGTATTA	AAAACCTCTT	TTATATATAA	GTCCGCTTAA	GGTGTTACTT
4501	CAATAATAAG	ATTAAAAATAG	CTTGCCCCCG	TTGCAGCGAT	GGGTATTTTT
	GTTATTATTC	TAATTTTATC	GAACGGGGGC	AACGTCGCTA	CCCATAAAAA
4551	TCTAGTAAAA	TAAAAGATAA	ACTTAGACTC	AAAACATTTA	CAAAAACAAC
	AGATCATTTT	ATTTTCTATT	TGAATCTGAG	TTTTGTAAAT	GTTTTTGTTG
4601	CCCTAAAGTC	CTAAAGCCCA	AAGTGCTATG	CACGATCCAT	AGCAAGCCCA
	GGGATTTTCA	GATTTTCGGT	TTCACGATAC	GTGCTAGGTA	TCGTTTCGGT
4651	GCCCAACCCA	ACCCAACCCA	ACCCACCCCA	GTGCAGCCAA	CTGGCAAATA
	CGGGTTGGGT	TGGGTGGGGT	TGGGTGGGGT	CACGTCGGTT	GACCGTTTAT
4701	GTCTCCACCC	CCGGCACAT	CACCGTGAGT	TGTCCGCACC	ACCGCACGTC
	CAGAGGTGGG	GGCCGTGATA	GTGGCACTCA	ACAGGCGTGG	TGGCGTGCAG
4751	TCGCAGCCAA	AAAAAAAAAA	AGAAAGAAAA	AAAAGAAAAA	GAAAAACAGC
	AGCGTCGGTT	TTTTTTTTTT	TCTTCTTTTT	TTTTCTTTTT	CTTTTTGTCTG
4801	AGGTGGGTCC	GGGTCGTGGG	GGCCGGAAAA	GCGAGGAGGA	TCGCGAGCAG
	TCCACCCAGG	CCCAGCACCC	CCGGCCTTTT	CGCTCCTCCT	AGCGCTCGTC
4851	CGACGAGGCC	CGGCCCTCCC	TCCGCTTCCA	AAGAAACGCC	CCCCATCGCC
	GCTGCTCCGG	GCCGGGAGGG	AGGCGAAGGT	TTCTTTGCGG	GGGGTAGCGG
4901	ACTATATACA	TACCCCCCCC	TCTCCTCCCA	TCCCCCAAC	CCTACCACCA
	TGATATATGT	ATGGGGGGGG	AGAGGAGGGT	AGGGGGGTTG	GGATGGTGGT
4951	CCACCACCAC	CACCTCCTCC	CCCCTCGCTG	CCGGACGACG	AGCTCCTCCC
	GGTGGTGGTG	GTGGAGGAGG	GGGGAGCGAC	GGCCTGCTGC	TCGAGGAGGG
5001	CCCTCCCCCT	CCGCCGCCGC	CGGTAACCAC	CCCGCCCCTC	TCCTCTTTCT
	GGGAGGGGGA	GGCGGCGGCG	GCCATTGGTG	GGGCGGGGAG	AGGAGAAAGA

FIG._40G

82 / 154

5051 TTCTCCGTTT TTTTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT  
AAGAGGCAAA AAAAAAAGCA GAGCCAGAGC TAGAAACCGG AACCATCAAA

5101 GGGTGGGCGA GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG  
CCCACCCGCT CTCGCCGAAG CAGCGGGTCT AGCCACGCGC CCTCCCCGCC

BamHI

~~~~~

5151 GATCTCGCGG CTGGCGTCTC CGGGCGTGAG TCGGCCCGGA TCCTCGCGGG
CTAGAGCGCC GACCGCAGAG GCCCGCACTC AGCCGGGCCT AGGAGCGCCC

5201 GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG
CTTACCCCGA GAGCCTACAT CTAGAAGAAA GAAAGAAGAA AAACACCATC

5251 AATTTGAATC CCTCAGCATT GTTCATCGGT AGTTTTTCTT TTCATGATTT
TTAAACTTAG GGAGTCGTAA CAAGTAGCCA TCAAAAAGAA AAGTACTAAA

5301 GTGACAAATG CAGCCTCGTG CGGAGCTTTT TTGTAGC
CACTGTTTAC GTCGGAGCAC GCCTCGAAAA AACATCG

FIG. 40H

83 / 154

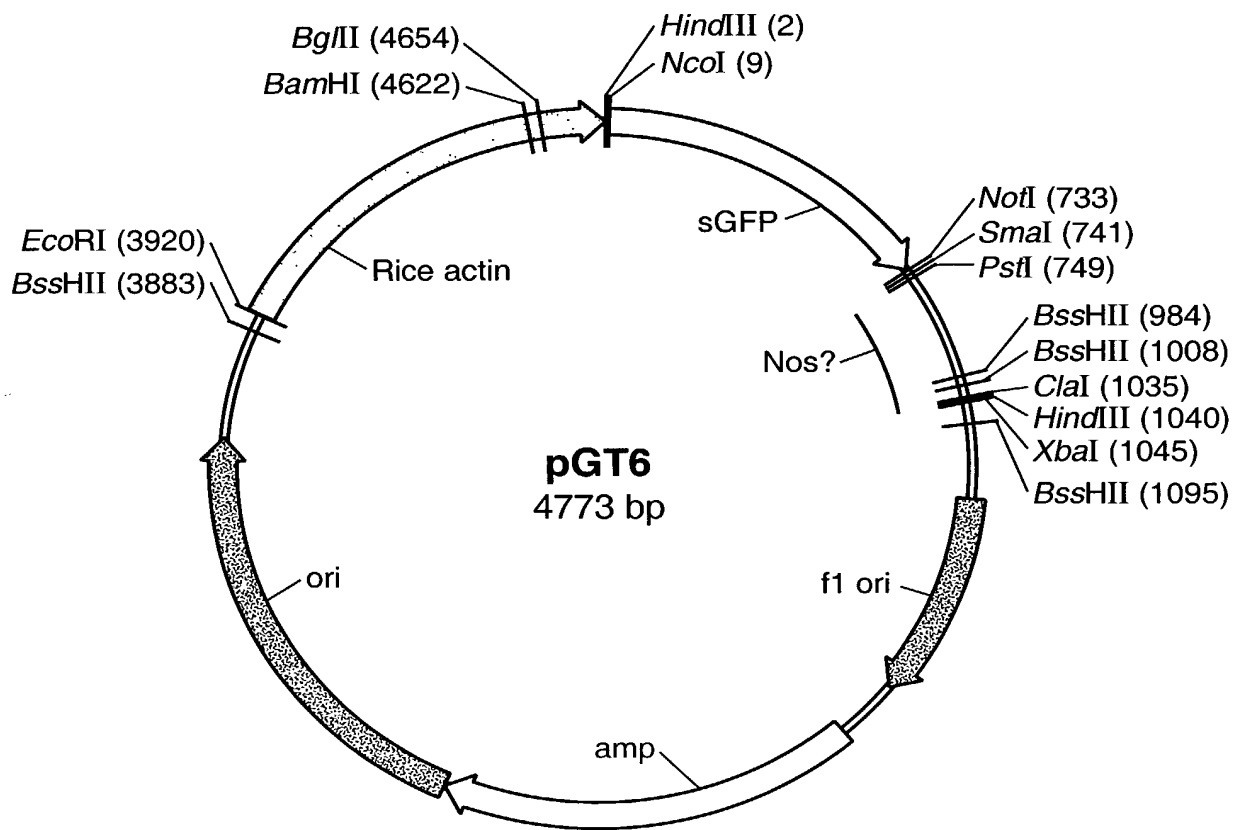


FIG.\_41A

[illegible]

FIG. 41B

```

SmaI
~~~~~
NotI      PstI
~~~~~
701 CTCACGGCAT GGACGAGCTG TACAAAGTAA GCGGCCGCC GGGCTGCAGG GAAACCACTG AAGGATGAGC
    GAGTGCCGTA CCTGCTCGAC ATGTTTCATTT CGCCGGCGGG CCCGACGTCC CTTTGGTGAC TTCCTACTCG

771 TGTAAAGAAG CAGATCGTTC AAACATTTGG CAATAAAGTT TCTTAAGATT GAATCCTGTT GCCGGTCTTG
    ACATTTCTTC GTCTAGCAAG TTTGTAAACC GTTATTTCAA AGAATTCTAA CTTAGGACAA CGGCCAGAAC

841 CGATGATTAT CATATAAATTT CTGTTGAATT ACGTTAAGCA TGTAAATAATT AACATGTAAT GCATGACGTT
    GCTACTAATA GTATATTAAA GACAACTTAA TGCAATTCTG ACATTATTAA TTGTACATTA CGTACTGCAA

911 ATTTATGAGA TGGGTTTTTA TGATTAGAGT CCCGCAATTA TACATTTAAT ACGCGATAGA AAACAAAATA
    TAAATACTCT ACCCAAAAAT ACTAATCTCA GGGCGTTAAT ATGTAAATTA TCGGCTATCT TTTGTTTAT

                                XbaI
                                ~~~~~
                                ClaI HindIII
                                ~~~~~
981 TAGCGCGCAA ACTAGGATAA ATTATCGCGC GCGGTGTCAAT CTATGTTACT AGATCGATAA GCTTCTAGAG
    ATCGCGCGTT TGATCCTATT TAATAGCGCG CGCCACAGTA GATACAATGA TCTAGCTATT CGAAGATCTC

                                BssHII
                                ~~~~~
                                BssHII
                                ~~~~~
1051 CGGCCGGTGG AGCTCCAATT CGCCCTATAG TGAGTCGTAT TACGCGCGCT CACTGGCCGT CGTTTACAA
    GCCGGCCACC TCGAGGTTAA GCGGGATATC ACTCAGCATA ATGCGCGCGA GTGACCCGCA GCAAAATGTT

1121 CGTCGTGACT GGGAAAACCC TGGCGTTACC CAACTTAATC GCCTTGACG ACATCCCCCT TTCGCCAGCT
    GCAGCACTGA CCTTTTGGG ACCGCAATGG GTTGAATTAG CGGAACGTCTG TGTAGGGGA AAGCGGTCGA

1191 GGCCTAATAG CGAAGAGGCC CGCACCGATC GCCCTTCCA ACAGTTGCGC AGCCTGAATG GCGAATGGGA
    CCGCATTATC GCTTCTCCGG GCGTGGCTAG CGGGAAGGGT TGTCAACGCG TCGGACTTAC CGCTTACCCT

1261 CGGCCCCTGT AGCGGCGCAT TAAGCGCGG GGGTGTGGTG GTTACGCGCA GCGTGACCGC TACACTTGCC
    GCGCGGGACA TCGCCGCGTA ATTGCGCGG CCCACACCAC CAATGCGCGT CGCACTGGCG ATGTGAACGG
```

FIG.\_41C

86 / 154

1331 AGCGCCCTAG CGCCCGCTCC TTTTCGCTTTC TTCCCTTCCT TTCTCGCCAC GTTCGCCGGC TTTCCCCGTC
TCGCGGGATC GCGGGCGAGG AAAGCGAAAG AAGGGAAGGA AAGAGCGGTG CAAGCGGCCG AAAGGGGCAG

1401 AAGCTCTAAA TCGGGGGCTC CCTTTAGGGT TCCGATTAG TGCTTTACGG CACCTCGACC CCAAAAAACT
TTCGAGATTT AGCCCCCGAG GGAATCCCA AGGCTAAATC ACGAAATGCC GTGGAGCTGG GGTTTTGTGA

1471 TGATTAGGT GATGGTTCAC GTAGTGGGCC ATCGCCCTGA TAGACGGTTT TTCCGCCCTTT GACGTTGGAG
ACTAATCCCA CTACCAAGTG CATCACCCGG TAGCGGACT ATCTGCCAAA AAGCGGAAA CTGCAACCTC

1541 TCCACGTTCT TTAATAGTGG ACTCTTGTTC CAAACTGGAA CAACACTCAA CCCTATCTCG GTCTATTCTT
AGTGCAAGA AATTATCAC TGAGAACAAG GTTTGACCTT GTTGTGAGTT GGGATAGAGC CAGATAAGAA

1611 TTGATTATA AGGATTTTG CCGATTTCCG CCTATTGGTT AAAAAATGAG CTGATTTAAC AAAAAATTAA
AACTAAATAT TCCCTAAAAC GGCTAAAGCC GGATAACCAA TTTTCTACTC GACTAAATTG TTTTAAATT

1681 CGCGAATTTT AACAAAATAT TAACGCTTAC AATTAGGTG GCACTTTTCG GGGAAATGTG CGCGGAACCC
GCGCTTAAA TTGTTTTATA ATTGCGAATG TTAATCCAC CGTGAAAAGC CCTTTTACAC GCGCCTTGGG

1751 CTATTGTTT ATTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCCT GATAAATGCT
GATAAACAAA TAAAAAGATT TATGTAAGTT TATACATAG CGAGTACTCT GTTATTGGGA CTATTACGA

1821 TCAATAATAT TGA AAAAGGA AGAGTATGAG TATTCACAT TTCCGTGTCTG CCTTATTCC CTTTTTTGCG
AGTTATTATA ACTTTTTCCT TCTCATACTC ATAAGTTGTA AAGGCACAGC GGGAAATAAGG GAAAAACGC

1891 GCATTTTGGC TTCTGTGTTT TGCTCACCCA GAAACGCTGG TGAAGATAA AGATGCTGAA GATCAGTTGG
CGTAAAACGG AAGGACAAA ACGAGTGGGT CTTTGCAGC ACTTTCATTT TCTACGACTT CTAGTCAACC

1961 GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTC GCCCCGAAGA
CACGTGCTCA CCCAATGTAG CTTGACCTAG AGTTGTCTGCC ATTCTAGGAA CTCTCAAAAG CCGGGCTTCT

2031 ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGTATGT GCGCGGTAT TATCCCGTAT TGACGCCGGG
TGCAAAAGGT TACTACTCGT GAAAATTCA AGACGATACA CCGCGCCATA ATAGGCATA ACTGCGGCC

2101 CAAGAGCAAC TCGGTCGCC CATACACTAT TCTCAGAATG ACTTGGTTGA GTACTACCA GTACAGAAA
GTTCTCGTTG AGCCAGCGGC GTATGTGATA AGAGTCTTAC TGAACCAACT CATGAGTGGT CAGTGTCTT

FIG. 41D

87 / 154

| | | | | | | | |
|------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 2171 | AGCATCTTAC | GGATGGCATG | ACAGTAAGAG | AATTATGCAG | TGCTGCCATA | ACCATGAGTG | ATAACACTGC |
| | TCGTAGAATG | CCTACCGTAC | TGTCATTCTC | TTAATACGTC | ACGACGGTAT | TGGTACTCAC | TATTGTGACG |
| 2241 | GGCCAACTTA | CTTCTGACAA | CGATCGGAGG | ACCGAAGGAG | CTAACCCTTT | TTTTGCACAA | CATGGGGGAT |
| | CCGGTTGAAT | GAAGACTGTT | GCTAGCCTCC | TGGCTTCCTC | GATTGGCGAA | AAAACGTGTT | GTACCCCCCTA |
| 2311 | CATGTAACTC | GCCTTGATCG | TTGGGAACCG | GAGCTGAATG | AAGCCATACC | AAACGACGAG | CGTGACACCA |
| | GTACATTGAG | CGGAACCTAGC | AACCCCTGGC | CTCGACTTAC | TTCGGTATGG | TTTGTCTGCTC | GCACTGTGGT |
| 2381 | CGATGCCTGT | AGCAATGGCA | ACAACGTTGC | GCAAACTATT | AACCTGGCGAA | CTACTTACTC | TAGCTTCCCCG |
| | GCTACGGACA | TCGTTACCGT | TGTTGCAACG | CGTTTGATAA | TTGACCCGCTT | GATGAATGAG | ATCGAAGGGC |
| 2451 | GCAACAATTA | ATAGACTGGA | TGGAGGCGGA | TAAAGTTGCA | GGACCACTTC | TGCGCTCGGC | CCTTCCGGCT |
| | CGTTGTTAAT | TATCTGACCT | ACCTCCGCCT | ATTTCACAGT | CCTGGTGAAG | ACGCGAGCCG | GGAAGGCCCGA |
| 2521 | GGCTGGTTTA | TTGCTGATAA | ATCTGGAGCC | GGTGAGCGTG | GGTCTCGCG | TATCATTTGCA | GCACTGGGGC |
| | CCGACCATAAT | AACGACTATT | TAGACCTCGG | CCACTCGCAC | CCAGAGCGCC | ATAGTAACGT | CGTGACCCCCG |
| 2591 | CAGATGGTAA | GCCCTCCCGT | ATCGTAGTTA | TCTACACGAC | GGGGAGTCA | GCAACTATGG | ATGAACGAAA |
| | GTCTACCAAT | CGGGAGGGCA | TAGCATCAAT | AGATGTGCTG | CCCCTCAGTC | CGTTGATACC | TACTTGCTTTT |
| 2661 | TAGACAGATC | GCTGAGATAG | GTGCCCTCACT | GATTAAGCAT | TGGTAACTGT | CAGACCAAGT | TTACTCATAT |
| | ATCTGTCTAG | CGACTCTATC | CACGGAGTGA | CTAATTCGTA | ACCATTGACA | GTCTGGTTCA | AATGAGTATA |
| 2731 | ATACTTTAGA | TTGATTTAAA | ACTTCATTTT | TAATTTAAAA | GGATCTAGGT | GAAGATCCTT | TTTGATAAATC |
| | TATGAAATCT | AACATAAATT | TGAAGTAAAA | ATTAAATTTT | CCTAGATCCA | CTTCTAGGAA | AAACTATTAG |
| 2801 | TCAATGACCAA | AATCCCTTAA | CGTGAGTTT | CGTTCCTACTG | AGCGTCAGAC | CCCGTAGAAA | AGATCAAAAGG |
| | AGTACTGGTT | TTAGGGAATT | GCACTCAAAA | GCAAGGTGAC | TCGCAGTCTG | GGGCATCTTT | TCTAGTTTCC |
| 2871 | ATCTTCTTGA | GATCCTTTT | TTCTGCGCGT | AATCTGCTGC | TTGCAAAACAA | AAAAACCAAC | GCTACCAGCG |
| | TAGAAGAACT | CTAGGAAAAA | AAGACGCGCA | TTAGACGACG | AACGTTTGTT | TTTTTTGGTGG | CGATGGTCCG |
| 2941 | GTGGTTTGTT | TGCCGGATCA | AGAGCTACCA | ACTCTTTTTC | CGAAGGTAAC | TGGCTTCAGC | AGAGCGCAGA |
| | CACCAACAA | ACGGCCTAGT | TCTCGATGGT | TGAGAAAAAG | GCTTCCATTG | ACCGAAGTCG | TCTCGCGTCT |

FIG.\_41E

3011 TACCAAATAC TGTCCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTTAC
ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC TTGAGACATC GTGGCGGATG

3081 ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT TACCGGGTTG
TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG TCACCGCTAT TCAGCACAGA ATGGCCCAAC

3151 GACTCAAGAC GATAGTTACC GGATAAGCG CAGCGGTCCG GCTGAACGGG GGGTTCGTGC ACACAGCCCA
CTGAGTTCTG CTATCAATGG CCTATTCCGC GTCGCCAGCC CGACTTGCCC CCCAAGCACG TGTGTCGGGT

3221 GCTTGGAGCG AACGACCTAC ACCGAACCTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG CCACGCTTCC
CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGAT ACTCTTTCGC GGTGCGAAGG

3291 CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT
GCTTCCCTCT TTCCGCCCTGT CCATAGGCCA TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG CTCCTTCGAA

3361 CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCCGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT
GGTCCCCCTT TGCGGACCAT AGAAATATCA GGACAGCCCA AAGCGGTGGA GACTGAACCTC GCAGCTAAAA

3431 TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGGC GCCTTTTAC GGTTCCTTGGC
ACACTACGAG CAGTCCCCC GCCTCGGATA CCTTTTGGC GTCGTTGCGC CGGAAAAATG CCAAGGACCG

3501 CTTTGTGCTG CCTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCTCGATT CTGTGGATAA CCGTATTACC
GAAAACGACC GGAAACGAG TGTACAAGAA AGGACGCAAT AGGGGACTAA GACACCTATT GGCATAATGG

3571 GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG
CGGAAACTCA CTCGACTATG GCGAGCGGCG TCGGCTTGCT TCGCTCGGTC GCTCAGTCAC TCGCTCCTTC

3641 CGGAAGAGCG CCCAATACGC AAACCGCCTC TCCCGCGCGG TTGGCCGATT CATTAATGCA GCTGGCACGA
GCCTTCTCGC GGGTTATGCG TTTGGCGGAG AGGGCGCGC AACCGGCTAA GTAATTACGT CGACCGTGCT

3711 CAGGTTTCCC GACTGGAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA GTTAGCTCAC TCATTAGGCA
GTCCAAAGGG CTGACCTTTC GCCCGTCACT CGCGTTGCGT TAATTACACT CAATCGAGTG AGTAATCCGT

3781 CCCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATGTTGT GTGGAATTGT GAGCGGATAA CAATTTCACA
GGGTCCCGAA ATGTGAAATA CGAAGGCCGA GCATACAACA CACCTTAACA CTCGCCATT GTTAAAGTGT

FIG.\_41F

89 / 154

| | BssHII
~~~~~ | | ECORI
~~~ |
|------|--|--|--------------|
| 3851 | CAGGAAACAG CTATGACCAT GATTACGCCA AGCGCGCAAT TAACCTCTAC TAAAGGGAAC AAAAGCTGGA
GTCCCTTTGTC GATACTGGTA CTAATGCGGT TCGCGCGTTA ATTGGGAGTG ATTTCCCTTG TTTTTCGACCT | | |
| 3921 | ATTCCACAAT GAACAATAAT AAGATTAAAA TAGCTTGCCC CCGTTGCAGC GATGGGTATT TTTTCTAGTA
TAAGGTGTTA CTTGTTATTA TTCTAATTTT ATCGAACGGG GGCAACGTCG CTACCCATAA AAAAGATCAT | | |
| 3991 | AAATAAAAGA TAAACTTAGA CTCAAAAACAT TTACAAAAAC AACCCCTAAA GTCCTAAAGC CCAAAGTGCT
TTTATTTTCT ATTGGAATCT GAGTTTTGTA AATGTTTTTG TTGGGGATTI CAGGATTTCG GGTTCACGA | | |
| 4061 | ATGCACGATC CATAGCAAGC CCAGCCCAAC CCAACCCACC CCAACCCACC CCAGTGCAGC CAACTGGCAA
TACGTGCTAG GTATCGTTCTG GGTCTGGGTTG GGTGGGTGG GGTTCACGTCG GGTGACCGTT | | |
| 4131 | ATAGTCTCCA CCCCCGGCAC TATCACCGTG AGTTGTCCGC ACCACCGCAC GTCTCGCAGC CAAAAAAA
TATCAGAGGT GGGGGCCGTG ATAGTGGCAC TCAACAGGCG TGTTGGCGTG CAGAGCGTCG GTTTTTTTTT | | |
| 4201 | AAAAGAAAGA AAAAAAGAA AAAGAAAAAC AGCAGGTGGG TCCGGGTCTG TCCGGCCCGA AAAGCGAGGA
TTTTCTTTCT TTTTCTTTCT TTTCTTTTGT TCGTCCACCC AGGCCACGCA CCCCCGGCCT TTTTCGCTCCT | | |
| 4271 | GGATCGCGAG CAGCGACGAG GCCCGGCCCT CCCTCCGCTT CCAAAGAAAC GCCCCCCATC GCCACTATAT
CCTAGCGCTC GTCGCTGCTC CGGGCCGGGA GGGAGGCGAA GGTCTCTTTG CGGGGGGTAG CCGTGATATA | | |
| 4341 | ACATACCCCC CCCTCTCTC CCATCCCCC AACCTACCA CCACCACCAC CACCACCTCC TCCCCCTCG
TGTATGGGG GGGAGAGGAG GGTAGGGGG TTGGGATGGT GGTGGTGGTG GTGGTGGAGG AGGGGGGAGC | | |
| 4411 | CTGCCGGACG ACGAGCTCCT CCCCCCTCCC CCTCCGCCGC CGCCGGTAAC CACCCCGCCC CTCTCTCTT
GACGGCCTGC TGCTCGAGGA GGGGGGAGGG GGAGGCGGCG CGGCCCATTG GTGGGCGGG GAGAGGAGAA | | |
| 4481 | TCCTTCTCCG TTTTCTTTT CGTCTCGGTC TCGATCTTTG GCCTTGGTAG TTTGGGTGG CGAGAGCGGC
AGAAAGAGGC AAAAAAAA GCAGAGCCAG AGTAGAAAC CGGAACCATC AAACCCACCC GCTCTCGCCG | | |
| 4551 | TTTCGTCGCC AGATCGGTGC GCGGGAGGG CGGGATCTCG CGGCTGGCGT CTCCGGGCGT GAGTCGGCCC
AAGCAGCGGG TCTAGCCACG CGCCCTCCCC GCCCTAGAGC GCCGACCGCA GAGGCCCGCA CTCAGCCGGG | | |

FIG. 41G

+

```

BamHI
~~~~~
4621 GGATCCTCGC GGGGAATGGG GCTCTCGGAT GTAGATCTTC TTTCTTTCTT CTTTTGTGG TAGAATTGA
    CCTAGGAGCG CCCCTTACCC CGAGAGCCTA CATCTAGAAG AAAGAAAGAA GAAAAACACC ATCTTAAACT

4691 ATCCCTCAGC ATTGTTTCATC GGTAGTTTTT CTTTTCATGA TTGTGACAA ATGCAGCCTC GTGCGGAGCT
    TAGGGAGTCG TAACAAGTAG CCATCAAAAA GAAAAGTACT AAACACTGTT TACGTCCGAG CACGCCCTCGA

4761 TTTTGTAGG TAG
    AAAACATCC ATC

BglII
~~~~~
    TTTCTTTCTT CTTTTGTGG TAGAATTGA
    GAAAAACACC ATCTTAAACT
    ATGCAGCCTC GTGCGGAGCT
    TACGTCCGAG CACGCCCTCGA
```

FIG.\_41H

+

91 / 154

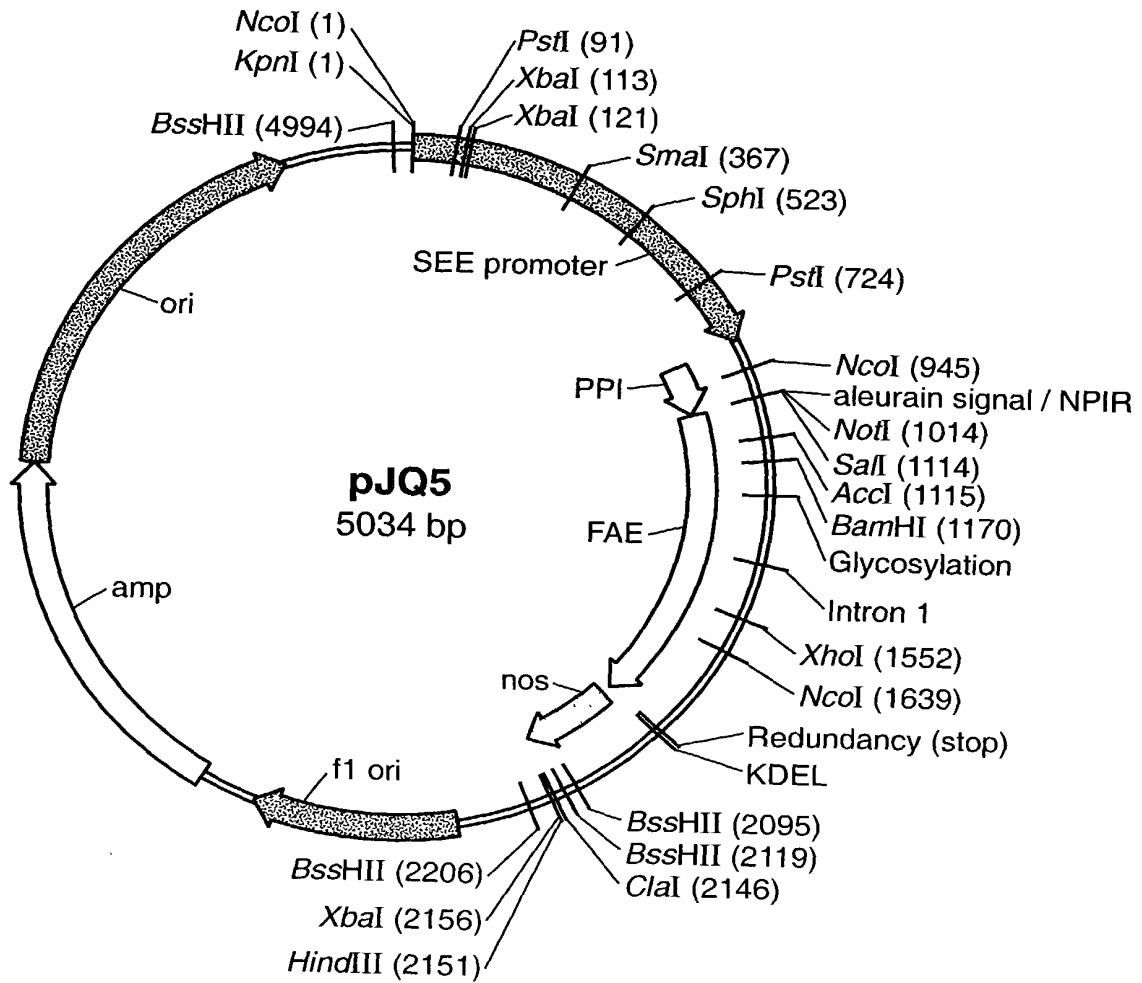


FIG. 42A

92 / 154

```

NcoI
~~~~~
KpnI
~
1  CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
   GTACCCCGGTC CATATTAATA CCCTATAGAG TTCGTTTATT AGCTTTATAG TGGTAACCGA TGTATATAGAC

                                     PstI
                                     ~~~~~
71  AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTGTG ATCTAGAACT CTAGATAGCA CAGCCACAGC
   TCGAGGCTCA AGACTGACGT CAGACCTACT GCGCACAAAC TAGATCTTGA GATCTATCGT GTCGGTGTCTG

                                     XbaI
                                     ~~~~~
141 ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTG GAGACGGAGC TCCTTCCTAC CTCCTGACGT
   TGGATGTCCT CACGCTGTGA ACACCTGACA TCATCACAAAC CTCCTGCCTCG AGAAAGGATG GAGGACTGCA

211 TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC CCAACAAAAA ATCGTCCCCC
   ACGGCGGCAA CAGGTAAGGT TGCCGTAGTG AGAGTTGGTT AGTGCGCGAG GGTGTGTTTA TAGCAGGGGG

281 ATGTCCTTGGC GGAGAGAGAG TACATACATG CTGTGCGGCC GTTTTGTGCT GAATCTCGCT TCCACTGGCC
   TACAGAAACCG CCTCTCTCTC ATGTATGTAC GACAGCGCGG CAAAAACAGA CTTAGAGCGA AGTGACCCGG

                                     SmaI
                                     ~~~~~
351 AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCTG TCACCCCTGG CGTCATGGGA
   TTAGTCGAGT CGAGGGCCCT CGAGTGAGTA AGTTCTAGGG TAGCAGCAGC AGTGGGGACC GCAGTACCCCT

421 TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAACCCCAAT
   ACCTTTTCTT GGAGGCAACG AGCCTACTCA GTCGGTATAG GGGCTTGCTT CATGACGTTT TATTGGGTTA

                                     SphI
                                     ~~~~~
491 TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTGTTGG CTTAATTGAC TTTATTTTTC
   AGTCTAAGGG GGTATATCTT TTCAATATCGT ACGAAAGCCC AAAACAAACC GAATTAACTG AAATAAAAAA

561 TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT CGAGACGGAT AATAGGCTGG
   AACCTCAACT TACGACTAAA CAACACATTT TACGGGTTGG TAGACTTATA GCTCTGCCCTA TTATCCGACC

```

FIG.. 42B

93 / 154

```

631  CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGGCATTACA GCTGGAGGCT
      GATTAATTAA ATATCGTTCT AAGACATCAC GTGTAGCGTT TATAGAAAGA CCCGTAATGT CGACCTCCGA

      PstI
      ~~~~~
701  TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG CGATGAGATG GGTATAAAAC
      AGTAGTCGGA CTTTGTGAGA CGTCTCGGAC TTCGTTCAAC ACTTCGCACC GCTACTCTAC CCATATTTTG

771  CCCCAGCACC GGGACGCGAG CTCCCAGCCTA CCAGTACCAT CTCGCCCTGC TCCCCCTGCC GGACGACCCA
      GGGGCCGTGG CCTGCGCTC GAGGGCGGAT GGTCATGGTA GAGCGGAGCG AGGGGACGG CCTGCTGGGT

841  GTAAAATACT GTTGCCCACT CGCCGGCGAG ATGGMCGTGC ACAAGGAGGT SAACTTCGTS GCCTACCTCC
      CATTTTATGA CAACGGGTGA GCGGCCGCTC TACCKGCACG TGTTCCTCCA STTGAAGCAS CGGATGGAGG

      NcoI
      ~~~~~
911  TGATCGTSTC CGGCTCCTC TTGCTCGTST CCGCCATGGA GCACGTGGAC GCCAAGGCCT GCACCKKCGA
      ACTAGCASA GCCGGAGGAG AACGAGCASA GCGGTACCT CGTGACCTG CCGTTCGGA CGTGGMGCT

      NotI
      ~~~~~
981  GTGCGGCAAC CTCGGCTTCG GCATCTGCC CCGGCCCGCC TCCACGCAGG GCATCTCCGA AGACCTCTAC
      CACGCCGTTG GAGCCGAAGC CGTAGACGG CGCCCGCGG AGGTGCGTCC CGTAGAGGCT TCTGGAGATG

      SalI
      ~~~~~
      AccI
      ~~~~~
1051 AGCCGTTTAG TCGAAATGGC CACTATCTCC CAAGCTGCCT ACGCCGACCT GTGCAACATT CCGTCGACTA
      TCGGCAATC AGCTTTACCG GTGATAGAGG GTTCGACGGA TCGCGCTGGA CACGTTGTAA GGCAGCTGAT

      BamHI
      ~~~~~
1121 TTATCAAGGG AGAGAAAATT TACAATTCTC AAAGTGACAT TAACGGATGG ATCCTCCGCG ACGACAGCAG
      AATAGTTCCC TCTCTTTTAA ATGTTAAGAG TTTGACTGTA ATTGCCTACC TAGGAGGCGC TGCTGTCGTC
```

FIG.\_42C

94 / 154

```
1191 CAAAGAAATA ATCACCGTCT TCCGTGGCAC TGGTAGTGAT ACGAATCTAC AACTCGATAC TAACTACACC
      GTTTCCTTAT TAGTGGCAGA AGGCACCGTG ACCATCACTA TGCTTAGATG TTGAGCTATG ATTGATGTGG

1261 CTCACGCCCT TCGACACCCCT ACCACAATGC AACGGTTGTG AAGTACACGG TGGATATTAT ATTGGATGGG
      GAGTGGCGAA AGCTGTGGGA TGGTGTACG TTGCCAACAC TTCATGTGCC ACCTATAATA TAACCTACCC

1331 TCTCCGTCCA GGACCAAGTC GAGTCGCTTG TCAAAACAGCA GGTAGCCAG TATCCGGACT ACGCGCTGAC
      AGAGGCAGGT CCTGGTTCAG CTCAGCGAAC AGTTTGTCTG CCAATCGGTC ATAGGCCCTGA TGCGCGACTG

1401 CGTGACCGGC CACKCCCTCG GCGCCTCCCT GCGGGCACTC ACTGCCGCC AGCTGTCTGC GACATACGAC
      GCACTGGCCG GTGMGGGAGC CGCGGAGGGA CCGCCGTGAG TGACGGCGGG TCGACAGACG CTGTATGCTG

1471 AACATCCGCC TGTACACCTT CGGCGAACC CGCAGCGGCA ATCAGGCCCT CGCGTCGTAC ATGAACGATG
      TTGTAGGCGG ACATGTGGAA GCCGCTTGGC GCGTCGCCGT TAGTCCGGAA GCGCAGCATG TACTTGCTAC

      XhoI
      ~~~~~

1541 CCTTCCAAGC CTCGAGCCCA GATACGACGC AGTATTTCG GGTCACTCAT GCCAACGACG GCATCCCAAA
      GGAAGGTTCT GAGCTCGGGT CTATGCTGCG TCATAAAGGC CCAGTGAGTA CGGTTGCTGC CGTAGGGTTT

      NcoI
      ~~~~~

1611 CCTGCCCCCG GTGGAGCAGG GGTACGCCCA TGGCGGTGTA GAGTACTGGA GCGTTGATCC TTACAGCGCC
      GGACGGGGGC CACCTCGTCC CCATGCGGGT ACCGCCACAT CTCATGACCT CGCAACTAGG AATGTCGCGG

1681 CAGAACACAT TTGCTGCAC TGGGGATGAA GTGCAGTGCT GTGAGGCCCA GGGCGGACAG GGTGTGAATA
      GTCTTGTA AACAGACGTG ACCCCTACTT CAGGTCACGA CACTCCGGGT CCCGCCCTGTC CCACACTTAT

1751 ATGCGCACAC GACTTATTTT GGGATGACGA GCGGAGCCTG TACATGGTGA TCAGTCATTT CAGCCTCCCC
      TACGCGGTGT CTGAATAAAA CCCTACTGCT CGCCTCGGAC ATGTACCACT AGTCAGTAAA GTCGGAGGGG

1821 GAGTGATCCA GGAAGATGG ATGTCCTGGA GAGGGGGCCG CGTAACCACT GAAGGATGAG CTGTAAAGAA
      CTCACATGGT CCTTCTTACC TACAGGACCT CTCCTCCGCT GCAATTGGTGA CTTCCCTACTC GACATTTCTT
```

FIG..42D

1209, 470202

95 / 154

```

1891 GCAGATCGTT CAAACATTG GCAATAAAGT TTCTTAAGAT TGAATCCTGT TGCCGGTCTT GCGATGATTA
    CGTCTAGCAA GTTGTAAAC CGTTATTCA AAGAATTCTA ACTTAGACA ACGGCCAGAA CGCTACTAAT

1961 TCATATAAAT TCTGTTGAAT TACGTTAAGC ATGTAATAAT TAACATGTAA TGCATGACGT TATTATGAG
    AGTATATTAA AGACAACCTA ATGCAATTCG TACATTATTA ATTGTACATT ACGTACTGCA ATAAATACTC

2031 ATGGGTTTTT ATGATTAGAG TCCCGCAATT ATACATTTAA TACGCGATAG AAAACAAAAT ATAGCGCGCA
    TACCCAAAAA TACTAATCTC AGGCGGTAA TATGTAAATT ATGCGCTATC TTTTGTTTTA TATCGCGCGT

                                     BssHII
                                     ~~~~~
                                     ClaI HindIII
                                     ~~~~~
                                     XbaI
                                     ~~~~~

2101 AACTAGGATA AATTATCGG CGCGGTGTCA TCTATGTTAC TAGATCGATA AGCTTCTAGA GCGGCCGGTG
    TTGATCCTAT TTAATAGCGC GCGCCACAGT AGATACAATG ATCTAGCTAT TCGAAGATCT CGCCGGCCAC

                                     BssHII
                                     ~~~~~

2171 GAGCTCCAAT TCGCCCTATA GTGAGTCGTA TTACGCGCGC TCACTGGCCG TCGTTTACA ACGTCGTGAC
    CTCGAGGTTA AGCGGGATAT CACTCAGCAT AATGCGCGCG AGTGACCGGC AGCAAAATGT TGCAGCACTG

2241 TGGGAAAACC CTGGCGTTAC CCAACTTAAT CGCCTTGCG CACATCCCCC TTTCGCCAGC TGGCGTAATA
    ACCCTTTTGG GACCGCAATG GGTGAATTA GCGGAACGTC GTGTAGGGG AAAGCGGTCG ACCGCATTAT

2311 GCGAAGAGGC CCGCACCGAT CGCCCTTCCC AACAGTTGCG CAGCCTGAAT GGCGAATGGG ACGCGCCCTG
    CGCTTCTCCG GCGGTGGCTA GCGGGAAGGG TTGTCAACGC GTCGGACTTA CCGCTTACCC TGC CGCGGGAC

2381 TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT GGTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA
    ATCGCCGCGT AATTCGCGCC GCCCACACCA CCAATGCGCG TCGCACTGGC GATGTGAACG GTCGCGGGAT

2451 GCGCCCGCTC CTTTCGCTTT CTTCCTTCC TTTCGCGCA CGTTCGCGG CTTTCCCGGT CAAGCTCTAA
    CGCGGGCGAG GAAAGCGAAA GAAGGAAGG AAAGAGCGGT GCAAGCGGCC GAAAGGGCA GTTCGAGATT

```

FIG..42E

96 / 154

2521 ATCGGGGGCT CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTAGGG
TAGCCCCCGA GGGAAATCCC AAGGCTAAAT CACGAAATGC CGTGGAGCTG GGGTTTTTTG AACATAATCCC

2591 TGATGGTTCA CGTAGTGGG CATCGCCCTG ATAGACGGTT TTTCGCCCTT TGACGTTGGA GTCCACGTTTC
ACTACCAAGT GCATCACCCG GTAGCGGGAC TATCTGCCAA AAGCGGGAA ACTGCAACCT CAGGTGCAAG

2661 TTTAATAGTG GACTCTTGTT CCAAACTGGA ACAACACTCA ACCCTATCTC GGTCTATTCT TTTGATTATAT
AAATTATCAC CTGAGAACAA GGTTTGACCT TGTGTGAGT TGGGATAGAG CCAGATAAGA AAATAAATA

2731 AAGGGATTTT GCCGATTTCG GCCTATTGGT TAAAAAATGA GCTGATTTAA CAAAAATTAA ACGCGAATTT
TTCCCTAAAA CGGCTAAAGC CGGATAACCA ATTTTTTACT CGACTAAAT GTTTTTAAAT TGCCTTAAA

2801 TAAACAAAATA TTAACGCTTA CAATTAGGT GGCACTTTTC GGGGAAATGT GCGCGGAACC CCTATTGTGT
ATTGTTTTAT AATTGCGAAT GTTAAATCCA CCGTGAAAAG CCCCTTTACA CGCGCCTTGG GGATAAACAA

2871 TATTTTCTA AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAATGC TTCAATAATA
ATAAAAAGAT TTATGTAAAT TTATACATAG GCGAGTACTC TGTATTGGG ACTATTTACG AAGTTATTAT

2941 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCGTGTC GCCCTTATTC CCTTTTTTGC GGCAATTTCG
AACTTTTTCC TTCTCATACT CATAAAGTTGT AAAGGCACAG CGGGAATAAG GGAATAAAGG CCGTAAAAAG

3011 CTTCCCTGTTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA AGATCAGTTG GGTGCACGAG
GAAGGACAAA AACGAGTGGG TCTTTGCGAC CACTTTCATT TTCTACGACT TCTAGTCAAC CCACGTGCTC

3081 TGGGTTACAT CGAACTGGAT CTCAACAGCG GTAAGATCCT TGAGAGTTT CGCCCCGAAG AACGTTTTC
ACCCAATGTA GCTTGACCTA GAGTTGTGCG CATCTAGGA ACTCTCAAA GCGGGGCTTC TTGCAAAAAGG

3151 AATGATGAGC ACTTTTAAAG TTCTGCTATG TGGCGCGGTA TTATCCCGTA TTGACGCCGG GCAAGAGCAA
TTACTACTCG TGAATAATTC AAGACGATAC ACCCGGCCAT AATAGGCAAT AACTGCGGCC CGTTCTCGTT

3221 CTCGGTCGCC GCATACACTA TTCTCAGAAAT GACTTGGTTG AGTACTCACC AGTCACAGAA AAGCATCTTA
GAGCCAGCGG CGTATGTGAT AAGAGTCTTA CTGAACCAAC TCATGAGTGG TCAGTGTCTT TTTCGTAGAAT

3291 CGGATGGCAT GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT GATAACACTG CGGCCAACTT
GCCTACCGTA CTGTCATTCT CTTAATACGT CACGACGGTA TTGGTACTCA CTATTGTGAC GCCGGTTGAA

FIG. 42F

97 / 154

3361 ACTTCTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTTCACA ACATGGGGA TCATGTAACT
TGAAGACTGT TGCTAGCCTC CTGGCTTCCT CGATTGGCGA AAAAACGTGT TGTACCCCTT AGTACATTGA
3431 CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA GCGTGACACC ACGATGCCCTG
GCGGAAC TAG CAACCCCTGG CCTCGACTTA CTTTCGGTATG GTTTGCTGCT CGCACTGTGG TGCTACGGAC
3501 TAGCAATGGC AACAAACGTTG CGCAAACTAT TAACTGGCGA ACTACTTACT CTAGCTTCCC GGCAACAAT
ATCGTTACCG TTGTTGCAAC GCGTTTGATA ATTGACCGCT TGATGAATGA GATCGAAGGG CCGTTGTTAA
3571 AATAGACTGG ATGGAGGCGG ATAAAGTTGC AGGACCACTT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT
TTATCTGACC TACCTCCGCC TATTTCAACG TCCTGGTGAA GACGCGAGCC GGAAGGCCG ACCGACCAAA
3641 ATTGCTGATA AATCTGGAGC CGGTGAGCGT GGGTCTCGCG GTATCATTCG AGCACTGGG CCAGATGGTA
TAACGACTAT TTAGACCTCG GCCACTCGCA CCCAGAGCGC CATAGTAACG TCGTGACCCC GGTCTACCAT
3711 AGCCCTCCCG TATCGTAGTT ATCTACACGA CGGGGAGTCA GGCAACTATG GATGAACGAA ATAGACAGAT
TCGGGAGGCG ATAGCATCAA TAGATGTGCT GCCCCTCAGT CCGTTGATAC CTACTTGCTT TATCTGTCTA
3781 CGCTGAGATA GGTGCCTCAC TGATTAAGCA TTGGTAACTG TCAGACCAAAG TTTACTCATA TATACTTTAG
GCGACTCTAT CCACGGAGTG ACTAATTCGT AACCATTGAC AGTCTGGTTC AAATGAGTAT ATATGAAATC
3851 ATTGATTTAA AACTTCATTT TTAATTTAAA AGGATCTAGG TGAAGATCCT TTTTGATAAT CTCATGACCA
TAACTAAATT TTGAAGTAAA AATTAATTT TCCTAGATCC ACTTCTAGGA AAAACTATTA GAGTACTGGT
3921 AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGCGTCAGA CCCCCTAGAA AAGATCAAAG GATCTTCTTG
TTTAGGGAAT TGCACTCAA AGCAAGGTGA CTCGCAGTCT GGGGCATCTT TTCTAGTTTC CTAGAAGAAC
3991 AGATCCCTTT TTTCTGCGCG TAATCTGCTG CTTGCAAAACA AAAAAACCAC CGCTACCAGC GGTGGTTTGT
TCTAGGAAA AAAGACGCGC ATTAGACGAC GAACGTTTGT TTTTITGGTG GCGATGGTCG CCACCAAAAC
4061 TTGCCGGATC AAGAGCTACC AACTCTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAATA
AACGGCCTAG TTCTCGATGG TTGAGAAAAA GGCTTCCATT GACCGAAGTC GTCTCGCGTC TATGGTTTAT
4131 CTGTCCTTCT AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA CATACCTCGC
GACAGGAAGA TCACATCGGC ATCAATCCGG TGGTGAAGTT CTTGAGACAT CGTGGCGGAT GTATGGAGCG

FIG. 42G

98 / 154

4201 TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGACTCAAGA
AGACGATTAG GACAAATGGTC ACCGACGACG GTCACCGCTA TTCAGCACAG AATGGCCCAA CCTGAGTTCT

4271 CGATAGTTAC CGGATAAGGC GCAGCGGTGC GGTGAACGG GGGGTTTCGTG CACACAGCCC AGCTTGGAGC
GCTATCAATG GCCTATTCCG CGTCGCCAGC CCGACTTGCC CCCCAGCAC GTGTGTCGGG TCGAACCTCG

4341 GAAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCT ATGAGAAAGC GCCACGCTTC CCGAAGGGAG
CTTGCTGGAT GTGGCTTGAC TCTATGGATG TCGCACTCGA TACTCTTTCG CGGTGCGAAG GGCTTCCCTC

4411 AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTCCGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA
TTTCCGCCTG TCCATAGGCC ATTGCGCCGTC CCAGCCTTGT CCTCTCGCGT GCTCCCTCGA AGGTCCCCCT

4481 AACGCCCTGGT ATCTTTATAG TCCTGTGCGG TTTCCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT
TTGCGGACCA TAGAAATATC AGGACAGCCC AAAGCGGTGG AGACTGAACT CGCAGCTAAA AACACTACGA

4551 CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCTTTTTA CGGTTCCTGG CCTTTTGCTG
GCAGTCCCCC CGCCTCGGAT ACCTTTTTCG GGTGCTTGCG CCGGAAAAAT GCCAAGGACC GGAAAAACGAC

4621 GCCTTTGTCT CACATGTTCT TTCTCTGCGTT ATCCCTTGAT TCTGTGGATA ACCGTATTAC CGCCTTTGAG
CGGAAAAACA GTGTACAAGA AAGGACGCAA TAGGGGACTA AGACACCTAT TGGCATAATG GCGGAAACTC

4691 TGAGCTGATA CCGCTCGCCG CAGCCGAACG ACCGAGCGCA GCGAGTCAGT GAGCGAGGAA GCGGAAGAGC
ACTCGACTAT GCGGAGCGGC GTCGGCTTGC TGGCTCGCGT CGCTCAGTCA CTCGCTCCTT CGCCTTCTCG

4761 GCCCAATACG CAAACCGCCT CTCGCCGCGC GTTGGCCGAT TCATTAATGC AGCTGGCACG ACAGGTTTCC
CGGGTTATGC GTTTGGCGGA GAGGGCGCG CAACCGGCTA AGTAATTACG TCGACCGTGC TGTCCAAAAGG

4831 CGACTGGAAA GCGGGCAGTG AGCGCAACGC AATTAATGTG AGTTAGCTCA CTCATTAGGC ACCCCAGGCT
GCTGACCTTT CGCCCGTCAC TCGCGTTGCG TTAATTACAC TCAATCGAGT GAGTAATCCG TGGGGTCCGA

4901 TTACACTTTA TGCTTCCGGC TCGTATGTTG TGTGGAATTG TGAGCGGATA ACAATTTAC ACAGGAAACA
AATGTGAAAT ACGAAGGCCG AGCATACAAC ACACCTTAAC ACTCGCCTAT TGTAAAGTG TGTCCTTTGT

FIG..42H

99 / 154

4971 GCTATGACCA TGATTACGCC AAGCGCGCAA TTAACCCCTCA CTAAAGGGAA CAAAAGCTGG GTAC
CGATACTGGT ACTAATGCGG TTCGCGCGTT AATTGGGAGT GATTCCCTT GTTTTCGACC CATG

BssHII
~~~~~

KpnI  
~~~~~

NcoI

FIG.\_42I

100 / 154

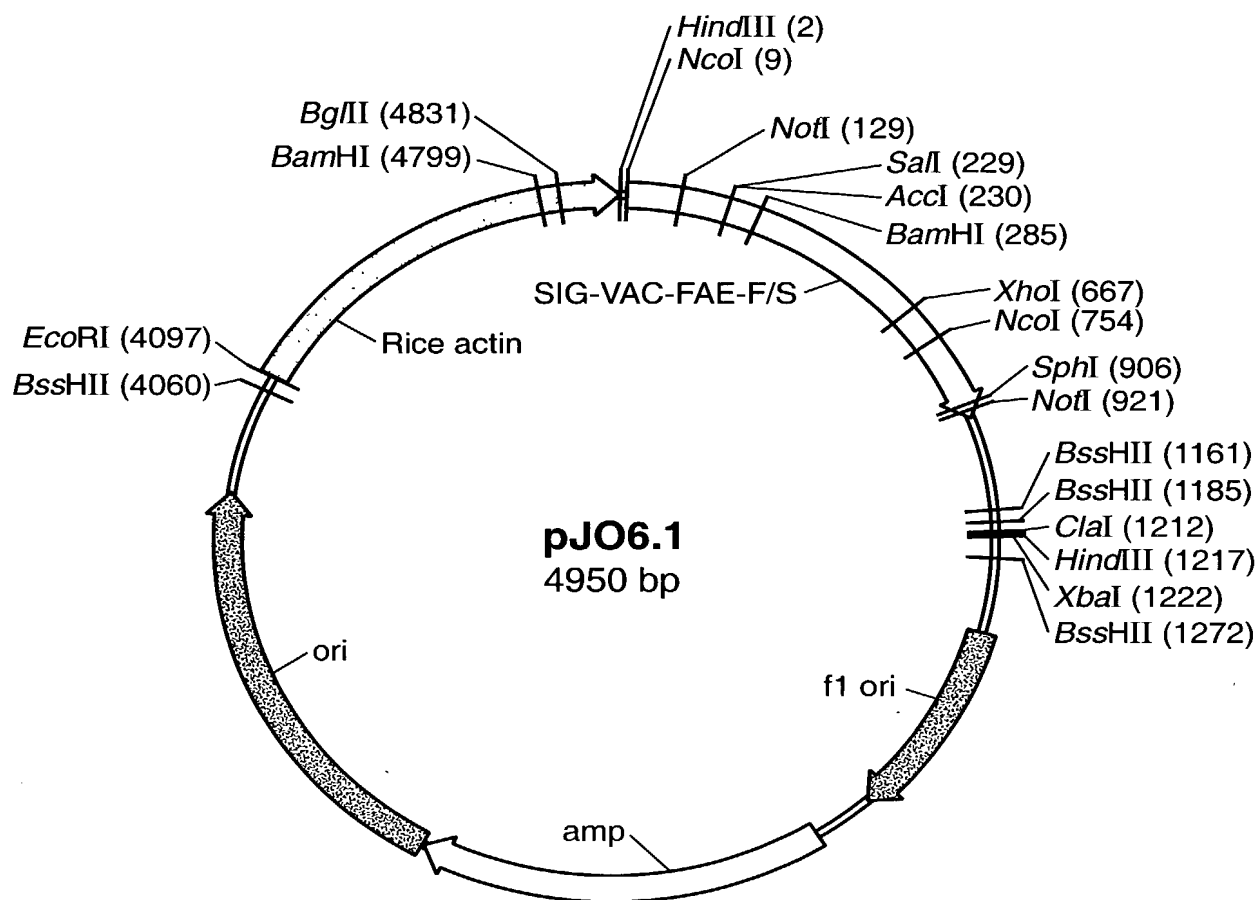


FIG. 43A

101 / 154

```
HindIII NcoI
~~~~~
1  AAGCTTACCA TGGCCACGC CCGCGTCTCTC CTCCTGGCGC TCGCCGTGCT GGCACGGCC GCGTCGCCG
   TTCGAATGGT ACCGGGTGG GCGCAGGAG GAGGACCGG GAGGCACGA CCGTGCCGG CGCAGCGGC

NotI
~~~~~
71  TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCGATCCG GCCCGTCACC GACCGCGCG GCGCCTCCAC
   AGCGGAGGAG GAGGAGGAAG CCGCTGAGGT TGGGCTAGGC CGGCGAGTG CTGGCGCGCC GCGGAGGTG

141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCCTACGCC
   CGTCCCCTAG AGGCTTCTGG AGATGTCGGC AAATCAGCTT TACCGGTGAT AGAGGGTTCTG ACGGATGCCG

SalI
~~~~~
AccI
~~~~~
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGAGAGA AAATTACAA TTCTCAAACT GACATTAACG
   CTGGACACGT TGTAAGGCAG CTGATAATAG TTCCCTCTCT TTTAAATGTT AAGAGTTTGA CTGTAATTGC

BamHI
~~~~~
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACGTGTA GTGATACGAA
   CTACCTAGGA GCGCGTGTG TCGTCGTTTC TTTATTAGTG GCAGAAAGCA CCGTGACCAT CACTATGCTT

351 TCTACAACCT GATACTAAT ACACCCCTAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
   AGATGTTGAG CTATGATTGA TGTGGGAGTG CGGAAAGCTG TGGGATGGTG TTACGTTGCC AACACTTCAT

421 CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAA CAGCAGGTTA
   GTGCCACCTA TAATATAACC TACCCAGAGG CAGGTCTTGG TTCAGCTCAG CGAACAGTTT GTCGTCCAAT

491 GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
   CCGTCATAGG CCTGATGCGC GACTGGCACT GGCCGGTGMG GGAGCCGCGG AGGACCCGCC GTGAGTGACG

561 CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
   GCGGGTCGAC AGACGCTGTA TGCTGTTGTA GCGGACATG TGAAGCCGC TTGGCGCGTC GCCGTTAGTC
```

FIG.\_43B

102 / 154

```

XhoI
~~~~~
631  GCCTTCGCGT  CGTACATGAA  CGATGCCCTTC  CAAGCCTCGA  GCCCAGATAC  GACGCAGTAT  TTCCGGGTCA
CGGAAGCGCA  GCATGTACTT  GCTACGGAAG  GTTCGGAGCT  CGGGTCTATG  CTGCGTCATA  AAGGCCCAGT

NcoI
~~~~~
701  CTCATGCCAA  CGACGGCATC  CCAAACCTGC  CCCCGGTGGA  GCAGGGGTAC  GCCCATGGCG  GTGTAGAGTA
GAGTACGGTT  GCTGCCGCTAG  GGTTTGGACG  GGGGCCACCT  CGTCCCCATG  CGGGTACCGC  CACATCTCAT

SphI
~~~~~
771  CTGGAGCGTT  GATCCTTACA  GCGCCCCAGAA  CACATTTGTC  TGCACCTGGG  ATGAAGTGCA  GTGCTGTGAG
GACCTCGCAA  CTAGGAATGT  CGCGGGTCTT  GTGTAAACAG  ACGTGACCCC  TACTTCACGT  CACGACACTC

NotI
~~~~~
841  GCCCAGGGCG  GACAGGGTGT  GAATAATGCG  CACACGACTT  ATTTTGGGAT  GACGAGCGGC  GCATGCACCT
CGGGTCCCCG  CTGTCCCCACA  CTTATTACGC  GTGTCTGAA  TAAAACCCCTA  CTGCTCGCCG  CGTACGTGGA

BssHI
~~~~~
911  GGCCGGTTCG  GGCCGCGGAA  ACCACTGAAG  GATGAGCTGT  AAAGAAGCAG  ATCGTTCAAA  CATTGGCAA
CCGGCCAGCG  CCGGCGCCTT  TGGTGACTTC  CTACTCGACA  TTTCTTCGTC  TAGCAAGTTT  GTAAACCGTT

BssHI
~~~~~
981  TAAAGTTTCT  TAAGATTGAA  TCCTGTGTGC  GGTCTTGCGA  TGATTATCAT  ATAATTTCTG  TTGAATTACG
ATTTCAAAGA  ATTCTAACTT  AGGACAACGG  CCAGAACGCT  ACTAATAGTA  TATTAAAGAC  AACTTAATGC

BssHI
~~~~~
1051  TTAAGCATGT  AATAATTAA  ATGTAATGCA  TGACGTTATT  TATGAGATGG  GTTTTTATGA  TTAGAGTCCC
AATTCGTACA  TTATTAATTG  TACATTACGT  ACTGCAATAA  ATACTCTACC  CAAAAATACT  AATCTCAGG

BssHI
~~~~~
1121  GCAATTATAC  ATTTAATACG  CGATAGAAAA  CAAAATATAG  CGCGCAAACT  AGGATAAATT  ATCGCGCGCG
CGTTAATATG  TAAATTATGC  GCTATCTTTT  GTTTTATATC  GCGCGTTTGA  TCCTATTATA  TAGCGCGCGC
```

FIG.-43C

103 / 154

```

XbaI
~~~~~
ClaI HindIII
~~~~~
1191 GTGTCATCTA TGTTACTAGA TCGATAAGCT TCTAGAGCGG CCGGTGGAGC TCCAATTGCG CCTATAGTGA
CACAGTAGAT ACAATGATCT AGCTATTCTGA AGATCTCGCC GGCCACCTCG AGGTTAAGCG GGATATCACT

BssHII
~~~~~
1261 GTCGTATTAC GCGCGCTCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCCTGG CGTTACCCAA
CAGCATAATG CCGCGGAGTG ACCGGCAGCA AAATGTTGCA GCACTGACCC TTTTGGGACC GCAATGGGTT

1331 CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC
GAATTAGCGG AACGTCGTGT AGGGGAAAG CCGTCGACCG CATTATCGCT TCTCCGGGCG TGGCTAGCGG

1401 CTTCCCAACA GTTGGCGAGC CTGAATGGCG AATGGGACGC GCCCTGTAGC GCGGCATTAAG GCGCGGCGGG
GAAAGGTTGT CAACGCGTCG GACTTACCGC TTACCCTGCG CCGGACATCG CCGCGTAATT CGCGCCGCCC

1471 TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTCTTTC
ACACCACCAA TCGCGCTCGC ACTGGCGATG TGAACGGTCG CGGGATCGCG GCGGAGGAAA GCGAAAGAAG

1541 CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCCTCAAG CTCATAATCG GGGGCTCCCT TTAGGGTTCC
GGAAGGAAAG AGCGGTGCAA GCGGCCGAAA GGGGCAGTTC GAGATTTAGC CCCCAGGGA AATCCCAAGG

1611 GATTTAGTGC TTTACGGCAC CTCGACCCCA AAAAAGTTGA TTAGGGTGAT GGTTCACGTA GTGGGCCATC
CTAAATCAG AATGCCGTG GAGCTGGGGT TTTTGAACCT AATCCCACTA CCAAGTGCAAT CACCCGGTAG

1681 GCCCTGATAG ACGGTTTTTC GCCCTTTGAC GTTGGAGTCC ACGTTCTTTA ATAGTGGACT CTTGTTCCAA
CGGGACTATC TGCCAAAAAG CCGGAAACTG CAACCTCAGG TGCAAGAAAT TATCACCTGA GAACAAGGTT

1751 ACTGGAACAA CACTCAACCC TATCTCGGTC TATTCCTTTG ATTTATAAGG GATTTTGCCG ATTTCCGGCTT
TGACCTTGTT GTGAGTTGGG ATAGAGCCAG ATAAGAAAAC TAAATATTCC CTAAACGGC TAAAGCCGGA

1821 ATTGGTTAAA AAATGAGCTG ATTTAACAAA AATTTAAGC GAATTTTAAAC AAAATATTAA CGCTTACAAT
TAACCAATTT TTTACTCGAC TAAATTGTTT TTAATTTGCG CTTAAAAATTG TTTTATAATT GCGAATGTTA
```

FIG. 43D

104 / 154

```
1891 TTAGGTGGCA CTTTTTCGGG AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA
AATCCACCGT GAAAAGCCCC TTACACGCG CCTTGGGAT AAACAAATAA AAAGATTAT GTAAAGTTTAT

1961 TGTATCCGCT CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
ACATAGGCGA GTACTCTGTT ATTGGGACTA TTTACGAAGT TATTATAACT TTTTCCTTCT CATACTCATA

2031 TCAACATTTT CGTGTGCGCC TTATTCCCTT TTTTGGGCA TTTTGCCTTC CTGTTTTTGC TCACCCAGAA
AGTTGTAAAG GCACAGCGGG AATAAGGGA AAAACGCCGT AAAACGGAAG GACAAAACG AGTGGGTCTT

2101 ACGCTGGTGA AAGTAAAGA TGCTGAAGAT CAGTTGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA
TGCAGCACT TTCATTTTCT ACGACTTCTA GTCAACCCAC GTGCTCACCC AATGTAGCTT GACCTAGAGT

2171 ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAAGC TTTTCCAATG ATGAGCACTT TTAAAGTTCT
TGTCGCCATT CTAGGAACTC TCAAAAGCGG GGCTTCTTGC AAAAGGTTAC TACTCGTGAA AATTCAAGA

2241 GCTATGTGGC GCGGTATTAT CCCGTATTGA CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACTATTCT
CGATACACCG CGCCATAATA GGCATAACT GCGGCCCGTT CTCGTTGAGC CAGCGCGTA TGTGATAAGA

2311 CAGAAATGACT TGGTTGAGTA CTCACCAATC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAAT
GTCTTACTGA ACCAACTCAT GAGTGGTCAG TGTCTTTTCG TAGAATGCCT ACCGTACTGT CATTCTCTTA

2381 TATGCAGTGC TGCCATAACC ATGAGTGATA ACAC TGCGGC CAAC TTACTT GACACAACGA TCGGAGGACC
ATACGTCACG ACGGTATTGG TACTCACTAT TGTGACGCG GTTGAATGAA GACTGTTGCT AGCCTCCTGG

2451 GAAGGAGCTA ACCGCTTTT TGCACAACAT GGGGATCAT GTAAC TCGC GTGATCGTTG GGAACCCGGAG
CTTCCCTCGAT TGGCGAAAAA ACGTGTGTA CCCCCTAGTA CATTGAGCGG AACTAGCAAC CCTTGGCCTC

2521 CTGAATGAAG CCATACCAA CGACGAGCGT GACACCACGA TGCCTGTAGC AATGGCAACA ACGTTGCGCA
GACTTACTTC GGTATGGTTT GCTGCTCGCA CTGTGGTGCT ACGGACATCG TTACCGTTGT TGCAACGCGT

2591 AACTATTAACT TGGCGAACTA CTTACTCTAG CTTCCCGCA ACAATTAATA GACTGGATGG AGCGGATAA
TTGATAATTG ACCGCTTGAT GAATGAGATC GAAGGCCCGT TGTTAATTAT CTGACCTACC TCCGCCATT

2661 AGTTGCAGGA CCACTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT
TCAACGTCCT GGTGAAGACG CGAGCCGGGA AGGCCGACCG ACCAATAAC GACTATTAG ACCTCGGCCA
```

FIG.\_43E

105 / 154

2731 GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAAGCC CTCCCCTATC GTAGTTATCT
CTCGCACCCA GAGCGCCATA GTAACGTCTGT GACCCCGGTC TACCATTCTGG GAGGGCATAG CATCAATAGA

2801 ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
TGTGCTGCCC CTCAGTCCGT TGATACCTAC TTGCTTTATC TGTCTAGCGA CTCTATCCAC GGAGTGACTA

2871 TAAGCATTTG TAACTGTCTAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAAACT TCATTTTTAA
ATTTCGTAACC ATTGACAGTC TGGTTCAAAAT GAGTATATAT GAAATCTAAC TAAATTTTGA AGTAAAAATT

2941 TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT CCCTTAACTG GAGTTTTCGT
AAATTTTCTT AGATCCACTT CTAGGAAAAA CTATTAGAGT ACTGGTTTTA GGGAAATTGCA CTCAAAAAGCA

3011 TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC TTCTTGAGAT CCTTTTTCG TGC GCGGTAAT
AGGTGACTCG CAGTCTGGGG CATCTTTTCT AGTTTCTTAG AAGAACTCTA GGAATAAAG AC GCGCATTA

3081 CTGCTGCTTG CAAACAAAAA AACCAACCGT ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT
GACGACGAAC GTTTGTTTTT TTGGTGGCGA TGGTCGCCAC CAAACAAACG GCCTAGTTCT CGATGGTTGA

3151 CTTTTTCCGA AGGTAACCTG CTTCAGCAGA GCGCAGATAC CAAATACTGT CCTTCTAGTG TAGCCGTTAGT
GAAAAAGGCT TCCATTGACC GAAGTCGTCT CGCGTCTATG GTTTATGACA GGAAGATCAC ATCGGCATCA

3221 TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
ATCCGGTGGT GAAGTTCTTG AGACATCGTG GCGGATGTAT GGAGCGAGAC GATTAGGACA ATGGTCACCG

3291 TGCTGCCAGT GCGGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG
ACGACGGTCA CCGCTATTCA GCACAGAAATG GCCCAACCTG AGTTCTGCTA TCAATGGCCT ATTCCGCGTC

3361 CGGTGCGGCT GAACGGGGG TTCTGTGCACA CAGCCAGCT TGGAGCGAAC GACCTACACC GAAC TGAGAT
GCCAGCCCCA CTTGCCCCCC AAGCACGTGT GTGCGGTGCA ACCTCGCTTG CTGGATGTGG CTTGACTCTA

3431 ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG GCGGACAGGT ATCCGGTTAAG
TGATGTGCGC ACTCGATACT CTTTCGCGGT GCGAAGGGCT TCCCTCTTTC CGCCTGTCCA TAGGCCATT

3501 CGGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGAAACG CCTGGTATCT TTATAGTCTCT
GCCGTCCCAG CCTTGTCTCTC TCGCGTGCTC CCTCGAAGGT CCCCTTTTGC GGACCATAGA AATATCAGGA

FIG.\_43F

106 / 154

```
3571 GTCGGGTTTC GCCACCTCTG ACTTGAGCGT CGATTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA
CAGCCCAAAG CCGTGGAGAC TGAACCTCGA GCTAAAAACA CTACGAGCAG TCCCCCGGCC TCGGATACCT

3641 AAAACGCCAG CAACGGGGCC TTTTACGGT TCCTGGCCCT TTTGCTGCACA TGTCTTTTCC
TTTTGCGGTC GTTGGGCCGG AAAAATGCCA AGGACCGGAA AACGACCGGA AACGAGTGT ACAAGAAAGG

3711 TGCGTTATCC CCGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC TCGCCGCAGC
ACGCAATAGG GGAATAAGAC ACCTATTGGC ATAATGGCGG AAATCACTC GACTATGGCG AGCGCGCTCG

3781 CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC AATACGCAAA CCGCCTCTCC
GCTTGCTGGC TCGCGTCGCT CAGTCACTCG CTCCCTCGCC TTCTCGCGGG TTATGCGTTT GCGGAGAGG

3851 CCGCGCGTTG GCCGATTCTAT TAATGCAGCT GGCACGACAG GTTCCCGGAC TGGAAAGCGG GCAGTGAGCG
GGCGCGCAAC CGGCTAAGTA ATTACGTCGA CCGTGTGTC CAAAGGCTG ACCTTTCGCC CGTCACTCGC

3921 CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT
GTTGCGTTAA TTACACTCAA TCGAGTGAGT AATCCGTGGG GTCCGAAATG TGAAATACGA AGGCCGAGCA

BssHII
~~~
3991 ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTACACACAG GAAACAGCTA TGACCATGAT TACGCCAAGC
TACAACACAC CTTAACACTC GCCTATTGTT AAAGTGTGTC CTTTGTGAT ACTGGTACTA ATGCGGTTTCG

BssHII
~~~~~
4061 GCGCAATTAA CCTCACTAA AGGGAACAA AGCTGGAATT CCACAATGAA CAATAATAAG ATTAAAAATAG
CGCGTTAATT GGGAGTGATT TCCCTTGTTT TCGACCTTAA GGTGTTACTT GTTATTATTC TAATTTTATC

4131 CTTGCCCCCG TTGCAGCGAT GGGTATTTT TCTAGTAAA TAAAGATAA ACTTAGACTC AAAACATTTA
GAACGGGGGC AACGTCGCTA CCCATAAAA AGATCATTTT ATTTCTATT TGAATCTGAG TTTTGTAAAT

4201 CAAAAACAAC CCTAAAGTC CTAAAGCCCA AAGTGCTATG CAGGATCCAT AGCAAGCCCA GCCCAACCCA
GTTTTTGTG GGGATTTCAG GATTTCGGGT TTCACGATAC GTGCTAGTA TCGTTCGGGT CGGGTTGGGT
```

FIG..43G

107 / 154

```
4271 ACCCAACCCA ACCCACCCCA GTGCAGCCAA CTGGCAAATA GTCTCCACCC CCGGCAC'TAT CACCGTGAGT
TGGGTTGGGT TGGGTGGGT CACGTCGGTT GACCGTTTAT CAGAGGTGGG GGCCGTGATA GTGGCACTCA

4341 TGTCCGCACC ACCGCACGTC TCGCAGCCAA AAAAAAATA AGAAGAAAA AAAAGAAAA GAAAAACAGC
ACAGGCGTGG TGGCGTGCAG AGCGTCGGTT TTTT'TTTT'TT TCTTCTTTT TTTTCTTTT CTTT'TTGTCG

4411 AGGTGGGTCC GGTTCGTGGG GGCCGGAAAA GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGGCCCTCCC
TCCACCCAGG CCCAGCACCC CCGGCCCTTT CGCTCCTCCT AGCGTCGTC GCTGCTCCGG GCCGGGAGGG

4481 TCCGCTTCCA AAGAAACGCC CCCCATCGCC ACTATATACA TACCCCCCCC TCTCCTCCCA TCCCCCAAC
AGGCGAAGGT TTCTTTTGGG GGGTAGCGG TGATATATGT ATGGGGGGG AGAGGAGGT AGGGGGGTG

4551 CCTACCACCA CCACCACCAC CACCTCCTCC CCCCTCGCTG CCGGACGACG AGCTCCTCCC CCCTCCCCCT
GGATGGTGGT GGTGGTGGTG GTGGAGGAGG GGGGAGCGAC GGCTGCTGC TCGAGGAGGG GGGAGGGGGA

4621 CCGCCGCCGC CGGTAACCA CCGCCCCCTC TCCTCTTTCT TTCTCCGTTT TTTT'TTTCGT CTCGGTCTCG
GGCGCGGCG GCCATTGGTG GGGCGGGGAG AGGAGAAAGA AAGAGGCAAA AAAAAAGCA GAGCCAGAGC

4691 ATCTTTGGCC TTGGTAGTTT GGTGGGGCGA GAGCGGCTTC GTCGCCCCAGA TCGGTGCGCG GGAGGGGCGG
TAGAAACCGG AACCATCAA CCCACCCGCT CTCGCCGAAG CAGCGGTCT AGCCACGCGC CCTCCCCGCC

BamHI BglII
~~~~~

4761 GATCTCGCG CTGGCGTCTC CGGGCGTGAG TCGGCCCCGA TCCTCGCGGG GAATGGGGCT CTCGGATGTA
CTAGAGCGCC GACCGCAGAG GCCCGCACTC AGCCGGGCCT AGGAGCGCCC CTTACCCCGA GAGCCTACAT

BglII
~~~~~

4831 GATCTCTTT CT'TCTTCTT TTTGTGGTAG AATTGAATC CCTCAGCAT GTTCATCGGT AGTTT'TCTT
CTAGAAGAAA GAAAGAAGAA AACACCATC TTAAACTTAG GGAGTCGTAA CAAGTAGCCA TCAAAAAGAA

4901 TTTCATGATT GTGACAAATG CAGCCTCGTG CGGAGCTTTT TTGTAGGTAG
AAGTACTAAA CACTGTTTAC GTCGGAGCAC GCCTCGAAAA AACATCCATC
```

FIG.\_43H

108 / 154

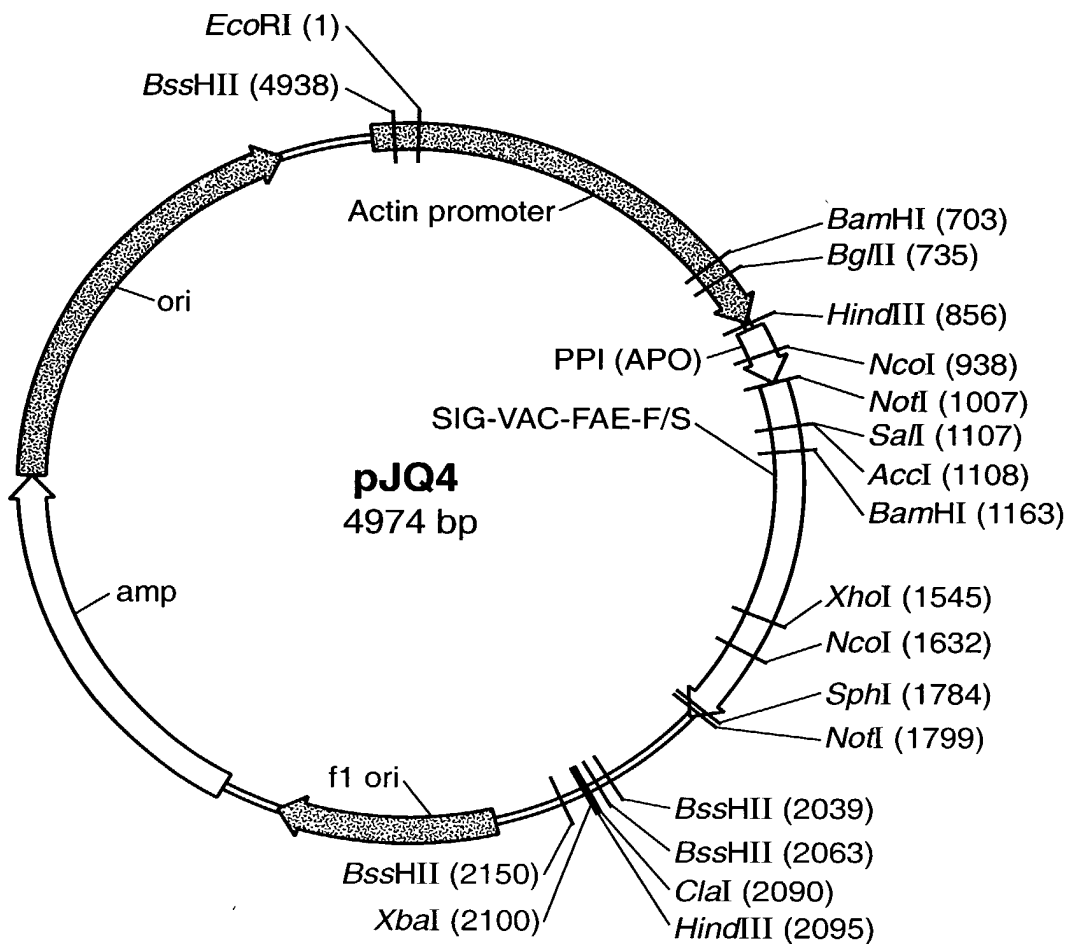


FIG. 44A

109 / 154

```
ECORI
~~~~~
1  AATTCCACAA TGAACAATAA TAAGATTAAA ATAGCTTGCC CCCGTTGCAG CGATGGGTAT TTTTCTAGT
   TTAAGGTGTT ACTTGTTATT ATCTAATTT TATCGAACGG GGGCAACGTC GCTACCCATA AAAAAGATCA

71  AAAATAAAAG ATAAACTTAG ACTCAAAAACA TTTACAAAAA CAACCCCTAA AGTCCCTAAAG CCCAAAGTGC
   TTTTATTTTC TATTGAATC TGAGTTTGT AAATGTTTTT GTTGGGATTC TCAGGATTC GGGTTTCACG

141 TATGCACGAT CCATAGCAAG CCCAGCCCAA CCCAACCCAC CCAGTGCAG CCAACTGGCA
   ATACGTGCTA GGTATCGTTC GGGTCGGGTT GGGTTGGGT GGTTCACGTC GGTGACCGT

211 AATAGTCTCC ACCCCGGCA CTATCACCGT GAGTTGTCCG CACCACCGCA CGTCTCGCAG CCAAAAAAAA
   TTATCAGAGG TGGGGCCCGT GATAGTGGCA CTCAACAGGC GTGGTGGCGT GCAGAGCGTC GGTTTTTTTT

281 AAAAAGAAAG AAAAAAAGA AAAAGAAAAA CAGCAGGTGG GTCCGGGTGC TGGGGGCCGG AAAAGCGAGG
   TTTTCTTTC TTTTCTTCT TTTTCTTCTT GTCGTCCACC CAGGCCCAGC ACCCCCGGCC TTTTCGCTCC

351 AGGATCGCGA GCAGCGACGA GGCCCGGCC TCCCTCCGCT TCCAAAGAAA CGCCCCCCAT CGCCACTATA
   TCCTAGCGCT CGTCGCTGCT CCGGGCCGG AGGAGGCGA AGGTTTCTTT GCGGGGGTA GCGGTGATAT

421 TACATACCCC CCCCTCTCCT CCATCCCCC CAACCTTACC ACCACACCA CCACCACCTC CTCCCCCCTC
   ATGTATGGGG GGGGAGAGGA GGTAGGGGG GTTGGGATGG TGGTGGTGGT GGTGGTGGAG GAGGGGGGAG

491 GCTGCCGGAC GACGAGCTCC TCCCCCTCC CCCTCCGCC CGCCCGGTAA CCACCCCGCC CCTCTCCTCT
   CGACGGCCTG CTGCTCGAG AGGGGGGAG GGGAGCGGC GCGGCCCAT GGTGGGCGCG GGAGAGGAGA

561 TTCTTTCTCC GTTTTCTTCT TCGTCTCGGT CTCGATCTTT GGCTTTGTA GTTTGGGTGG GCGAGAGCGG
   AAGAAAGAGG CAAAAAAA AGCAGAGCCA GAGTAGAAA CCGGAACCAT CAAACCCACC CGCTCTCGCC

631 CTTCGTCGCC CAGATCGGTG CGCGGGAGGG GCGGATCTC GCGGCTGGCG TCTCCGGGCG TGAGTCGGCC
   GAAGCAGCGG GTCTAGCCAC GCGCCCTCCC CGCCCTAGAG CGCCGACCGC AGAGGCCCGC ACTCAGCCCG

BamHI
~~~~~
701 CGGATCCTCG CGGGGAATGG GGCTCTCGGA TGATGATCTT CTTTCTTCT TCTTTTGTG GTAGAATTG
   GCCTAGGAGC GCCCCTTACC CCGAGAGCCT ACATCTAGAA GAAAGAAAGA AGAAAAACAC CATCTTAAAC

BglII
~~~~~
```

FIG. 44B

110 / 154

```
771 AATCCCTCAG CATGTTTCAT CGGTAGTTTT TCTTTTCATG ATTTGTGACA AATGCAGCCT CGTGCGGAGC
TTAGGGAGTC GTAACAAGTA GCCATCAAAA AGAAAAGTAC TAAACACTGT TTACGTCGGA GCACGCCCTCG

HindIII
~~~~~
841 TTTTTTGTAG GTAGAAGCTT ACMATGGMCG TGCACAAAGG GGTSAACTTC GTSGCCTACC TCCTGATCGT
AAAAAACATC CATCTTCGAA TGKTACCKGC ACGTGTTCTT CCASTTGAAG CASC GGATGG AGGACTAGCA

NcoI
~~~~~
911 SCTCGGCCTC CTC TTGCTCG TSTCCGCCAT GGAGCACGTG GACGCCAAGG CCTGCACCCCK CGAGTGCGGC
SGAGCCGGAG GAGAACGAGC ASAGGCGGTA CCTCGTGCAC CTGCGGTTCC GGACGTGGGM GCTCACGCCCG

NotI
~~~~~
981 AACCTCGGCT TCGGCATCTG CCGGCGGCC GCCTCCACGC AGGCATCTC CGAAGACCTC TACAGCCGTT
TTGGAGCCGA AGCCGTAGAC GGGCCGCCGG CGGAGGTGCG TCCCGTAGAG GCTTCTGGAG ATGTCGGCAA

Sali
~~~~~
AccI
~~~~~
1051 TAGTCGAAAT GGCCACTATC TCCCAAGCTG CCTACGCCGA CCTGTGCAAC ATTCCGTCGA CTATTATCAA
ATCAGCTTTA CCGGTGATAG AGGTTTCGAC GGATGCGGCT GGACACGTTG TAAGGCAGCT GATAATAGTT

BamHI
~~~~~
1121 GGGAGAGAAA ATTTACAATT CTCAAACTGA CATTAAACGA TGGATCCTCC GCGACGACAG CAGCAAAGAA
CCCTCTCTTT TAAATGTTAA GAGTTTGACT GTAATTGCCT ACCTAGGAGG CGCTGCTGTC GTCGTTTCTT

1191 ATAATCACCG TCTTCCGTGG CACTGGTAGT GATACGAATC TACAACTCGA TACTAACTAC ACCCTCACGC
TATTAGTGGC AGAAGGCACC GTGACCATCA CTATGCTTAG ATGTTGAGCT ATGATTGATG TGGGAGTGCG

1261 CTTTCGACAC CCTACCACAA TGCAACGGTT GTGAAGTACA CGGTGGATAT TATATTGGAT GGGTCTCCGT
GAAAGCTGTG GGATGGTGTT ACGTTGCCAA CACTTCATGT GCCACCTATA ATATAACCTA CCCAGAGGCA
```

FIG.\_44C

111 / 154

```
1331 CCAGGACCAA GTCGAGTCGC TTGTCAAACA GCAGGTTAGC CAGTATCCGG ACTACGCGCT GACCGTGACC
      GGTCCTGGTT CAGCTCAGCG AACAGTTTGT CGTCCAATCG GTCATAGGCC TGATGCGCGA CTGGCACTGG

1401 GGCCACKCCC TCGGCGCCTC CCTGGCGGCA CTCAC TGCCG CCCAGCTGTC TCGGACATAC GACAACATCC
      CCGGTGMGG AGCCGCGGAG GGACCGCCGT GAGTGACGGC GGTGCGACAG ACGCTGTATG CTGTTGTAGG

1471 GCCTGTACAC CTTCGGCGAA CCGCGCAGCG GCAATCAGGC CTTGCGGTGC TACATGAACG ATGCCCTTCCA
      CGGACATGTG GAAGCCGCTT GGCGCGTCGC CGTTAGTCCG GAAGCGCAGC ATGTACTTGC TACGGAAGGT

      XhoI
      ~~~~~

1541 AGCCTCGAGC CCAGATACGA CGCAGTATTT CCGGGTCACT CATGCCAACG ACGGCATCCC AAACCTGCCC
      TCGGAGCTCG GGTCATGCT GCGTCATAAA GGCCCAAGTGA GTACGGTTGC TGCCGTAGGG TTGGGACGGG

      NcoI
      ~~~~~

1611 CCGGTGGAGC AGGGGTACGC CCATGGCGGT GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCCAGAACA
      GGCCACCTCG TCCCCATGCG GGTACCGCCA CATCTCATGA CCTCGCAACT AGGAATGTGC CGGGTCTTGT

1681 CATTGTCTG CACTGGGGAT GAAGTGCAGT GCTGTGAGGC CCAGGGCGGA CAGGGTGTGA ATAATGCGCA
      GTAAACAGAC GTGACCCCTA CTTACAGTCA CGACACTCCG GTCCCCGCCT GTCCACACT TATTACGCGT

      SphI          NotI
      ~~~~~          ~~~~~

1751 CACGACTTAT TTTGGGATGA CGAGCGGCGC ATGCACCTGG CCGGTCGCGG CCGCGGAAAC CACTGAAGGA
      GTGCTGAATA AAACCTACT GCTCGCCGCG TACGTGGACC GGCCAGCGCC GGCGCCTTTG GTGACTTCCT

1821 TGAGCTGTAA AGAAGCAGAT CGTTCAAACA TTTGGCAATA AAGTTTCTTA AGATTGAATC CTGTTGCCGG
      ACTCGACATT TCTTCGTCTA GCAAGTTTGT AAACCGTTAT TTCAAAGAAAT TCTAACTTAG GACAAACGGCC

1891 TCTTGCATG ATTATCATAT AATTCTCTGT GAATTACGTT AAGCATGTAA TAAATTAACAT GTAATGCATG
      AGAACGCTAC TAATAGTATA TTAAAGACAA CTTAATGCAA TTCGTACATT ATTAATTGTA CATTACGTAC

1961 ACGTTATTTA TGAGATGGGT TTTTATGATT AGAGTCCCGC AATTATACAT TTAATACGCG ATAGAAAACA
      TGCAATAAAT ACTCTACCCA AAAATACTAA TCTCAGGGCG TTAATATGTA AATTATGCGC TATCTTTTGT
```

FIG..44D

2031 XbaI BssHII BssHII ClaI HindIII XbaI
         ~~~~~    ~~~~~    ~~~~~    ~~~~~    ~~~~~  
         AAATATAGCG CGCAAACTAG GATAAATTAT CGCGCGCGGT GTCATCTATG TTACTAGATC GATAAGCTTC  
         TTATATATCGC GCGTTTGATC CTATTTAATA GCGCGCGCCA CAGTAGATAC AATGATCTAG CTATTCGAAG  
         XbaI    BssHII    BssHII    BssHII    BssHII    BssHII  
         ~~~~~    ~~~~~    ~~~~~    ~~~~~    ~~~~~  
2101 TAGAGCGGCC GGTGGAGCTC CAATTCGCC TATAGTGAGT CGTATTACGC GCGCTCACTG GCCGTCGTTT
 ATCTCGCCCG CCACCTCGAG GTTAAGCGGG ATATCACTCA GCATAATGCG CGCGAGTGAC CGGCAGCAAA
2171 TACAACGTCG TGA CTGGGAA AACCTGGCG TTACCCAACT TAATCGCCTT GCAGCACATC CCCCTTTCGC
 ATGTTGCAGC ACTGACCCCTT TTGGGACCGC AATGGTTGA ATTAGCGGAA CGTCGTGTAG GGGGAAAGCG
2241 CAGCTGGCGT AATAGCGAAG AGGCCCGCAC CGATCGCCCT TCCCAACAGT TGCGCAGCCT GAATGGCGAA
 GTCGACCGCA TTATCGCTTC TCCGGGCGTG GCTAGCGGGA AGGTTGTCA ACGGTCGGA CTTACCGCTT
2311 TGGGACGCGC CCTGTAGCGG CGCATTAAGC GCGCGGGTG TGGTGGTTAC GCGCAGCGTG ACCGCTACAC
 ACCCTGCGCG GGACATCGCC GCGTAATTCG CGCCGCCAC ACCACCAATG CGCGTCGCAC TGGCGATGTG
2381 TTGCCAGCGC CCTAGCGCCC GCTCCTTTTCG CTTTCTTCCC TTCTCTTCTC GCCACGTTTCG CCGGCTTTCC
 AACGGTCGCG GGATCGCGGG CGAGGAAAGC GAAAGAAGG AAGGAAAGAG CCGTGCAAGC GGCCGAAAGG
2451 CCGTCAAGCT CTAAATCGGG GGCTCCCTTTT AGGGTTCCGA TTTAGTGCTT TACGGCACCT CGACCCCAA
 GGCAGTTCGA GATTTAGCCC CCGAGGGAAA TCCCAAGGCT AAATCACGAA ATGCCGTGGA GCTGGGGTTT
2521 AAACCTTGATT AGGGTGATGG TTCACGTAGT GGGCCATCGC CCTGATAGAC GGTTTTTCGC CCTTTGACGT
 TTTGAACATA TCCCACCTACC AAGTGCATCA CCGGTAGCG GGA CTATCTG CCAAAAAGCG GGAAACTGCA
2591 TGGAGTCCAC GTTCTTTAAT AGTGGACTCT TGTTCCAAAC TGGAACAACA CTCAACCCCTA TCCTCGGTCTA
 ACCTCAGGTG CAAGAAATTA TCACCTGAGA ACAAGTTTG ACCTTGTGTG GAGTTGGGAT AGAGCCAGAT
2661 TTCTTTTGAT TTATAAGGGA TTTTGCCGAT TTCGGCCCTAT TGGTTAAAAA ATGAGCTGAT TTAACAAAAA
 AAGAAAACTA AATATTCCCT AAAACGGCTA AAGCCGGATA ACCAATTTTT TACTCGACTA AATTGTTTTT

FIG..44E

113 / 154

2731 TTTAAACGCGA ATTTTAAACA AATATTAACG CTTACAATTT AGGTGGCACT TTTCGGGGAA ATGTGCGCGG
AAATGCGCT TAAAATGT TTTAATATGCG GAATGTTAAA TCCACCGTGA AAAGCCCCCTT TACACGCGCC

2801 AACCCCTATT TGTATTATTT TCTAAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA
TTGGGGATAA ACAAATAAAA AGATTATATG AAGTTATATC ATAGGCGAGT ACTCTGTTAT TGGGACTATT

2871 ATGCTTCAAT AATATTGA AAAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT
TACGAAGTTA TTATAACTTT TTCCCTTCTCA TACTCATAG TTGTAAAAGC ACAGCGGGAA TAAGGGAAAA

2941 TTGCGGCATT TTGCCCTTCT GTTTTGTCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA
AACGCCGTAA AACGGAAGGA CAAAACGAG TGGGTCTTTG CGACCACCTT CATTTCTAC GACTTCTAGT

3011 GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC
CAACCCACGT GCTCACCCAA TGTAGCTTGA CCTAGAGTTG TCGCCATTCT AGGAACTCTC AAAAGCGGGG

3081 GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTATTGACG
CTTCTTGCAA AAGTTACTA CTCGTGAAAA TTTCAAGACG ATACACCGCG CCATAATAGG GCATAACTGC

3151 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC
GGCCCGTTCT CGTTGAGCCA GCGGCGTATG TGATAAGAGT CTTACTGAAC CAACTCATGA GTGGTCAGTG

3221 AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC
TCTTTTCGTA GAATGCCCTAC CGTACTGTCA TTCTCTTAAT ACGTCACGAC GGTATTGGTA CTCACTATTG

3291 ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG
TGACGCCGGT TGAATGAAGA CTGTTGCTAG CCTCCTGGCT TCCTCGATTG GCGAAAAAAC GTGTTGTACC

3361 GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA
CCCTAGTACA TTGAGCGGAA CTAGCAACCC TTGGCCCTCGA CTTACTTCGG TATGGTTTGC TGCTCGCACT

3431 CACCACGATG CCTGTAGCAA TGGCAACAC GTTGCGCAAA CTATTAACTG GCGAACTACT TACTCTAGCT
GTGGTGTCTAC GGACATCGTT ACCGTTGTTG CAACGCGTTT GATAAATTGAC CGCTTGATGA ATGAGATCGA

3501 TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC
AGGCCCGTTG TTAATTATCT GACCTACCTC CGCCTATTTC AACGTCTTGG TGAAGACGCG AGCCGGGAAG

FIG..44F

114 / 154

3571 CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT
GCCGACCGAC CAAATAACGA CTATTTAGAC CTCGGCCACT CGCACCCAGA GCGCCATAGT AACGTCGTGA
3641 GGGGCCAGAT GGTAAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA
CCCCGGTCTA CCAATTCGGA GGGCATAGCA TCAATAGATG TGCTGCCCTT CAGTCCGTTG ATACCTACTT
3711 CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACGTGATTA AGCATTTGTA ACTGTCAGAC CAAGTTTACT
GCTTTATCTG TCTAGCGACT CTATCCACGG AGTGACTAAT TCGTAACCAT TGACAGTCTG GTTCAAAATGA
3781 CATATATACT TTAGATTGAT TTAAAACTTC ATTTTAAAT TAAAGGATC TAGGTGAAGA TCCTTTTGA
GTATATATGA AATCTAACTA AATTTTGAAG TAAAAATTAA ATTTTCCTAG ATCCACTTCT AGGAAAAACT
3851 TAAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT CAGACCCCGT AGAAAAAGATC
ATTAGAGTAC TGGTTTTAGG GAATTCACAT CAAAAGCAAG GTGACTCGCA GTCTGGGGCA TCCTTTCTAG
3921 AAAGGATCTT CTTGAGATCC TTTTTCCTG CGCGTAATCT GCTGCTTGCA AACAAAAAA CCACCGCTAC
TTTCCTAGAA GAACTCTAGG AAAAAAGAC GCGCATAGA CGACGAACGT TTGTTTTCCTT GGTGGCGATG
3991 CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACTGGCT TCAGCAGAGC
GTCGCCACCA AACAAACGGC CTAGTTCTCG ATGGTTGAGA AAAAGGCTTC CATTGACCGA AGTCGTCTCG
4061 GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTA GGCCACCACT TCAAGAACTC TGTAGCACCG
CGTCTATGGT TTATGACAGG AAGATCACAT CGGCATCAAT CCGGTGGTGA AGTTCTTGAG ACATCGTGGC
4131 CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCC TGCTTTACCG
GGATGTATGG AGCGAGACGA TTAGGACAAAT GGTCACCGAC GACGGTCACC GCTATTTCAGC ACAGAAATGGC
4201 GGTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGACACACA
CCAACTTGAG TTCTGCTATC AATGGCCTAT TCCGCGTCGC CAGCCCCGACT TGCCCCCCTAA GCACGTGTGT
4271 GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGCGCCACG
CGGGTCGAAC CTCGCTTGCT GGATGTGGCT TGACTCTATG GATGTCGCAC TCGATACTCT TTCGCGGTGC
4341 CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG
GAAGGGCTTC CCTCTTTCGG CCTGTCCATA GGCCATTTCG CGTCCAGCC TTGTCTCTCTC GCGTGTCTCC

FIG.\_44G

115 / 154

```
4411 AGCTTCCAGG GGAACAACGCC TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG
TCGAAGGTCC CCTTTGCGG ACCATAGAAA TATCAGGACA GCCCAAAGCG GTGGAGACTG AACTCGCAGC

4481 ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGAAA AACGCCAGCA ACGGGCCCTT TTTACGGTTC
TAAAAACACT ACGAGCAGTC CCCCCGCCTC GGATACCTTT TTGCGGTCGT TCGGCCGGAA AATGCCCCAAG

4551 CTGGCCCTTT GCTGGCCCTT TGCTCACATG TTCTTTTCCTG CGTTATCCCC TGATTCTGTG GATAAACCGTA
GACCGGAAAA CGACCGGAAA ACGAGTGTA AAGAAAGGAC GCAATAGGG ACTAAGACAC CTATTGGCAT

4621 TTACCGCCTT TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA
AATGGCGGAA ACTCACTCGA CTATGGCGAG CGGCGTCGGC TTGCTGGCTC GCGTCGCTCA GTCACTCGCT

4691 GGAAGCGGAA GAGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC CGATTCAATTA ATGCAGCTGG
CCTTCGCCCTT CTCGCGGGTT ATGCGTTTGG CGGAGAGGGG CGCGCAACCG GCTAAGTAAT TACGTCGACC

4761 CACGACAGGT TTCCCGACTG GAAAGCGGCG AGTGAGCGCA ACGCAATTA TGTGAGTTAG CTCACCTCAT
GTGCTGTCCA AAGGGCTGAC CTTTCGCCCG TCACCTGCGT TCGGTTAATT ACACTCAATC GAGTGAGTAA

4831 AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT
TCCGTGGGGT CCGAAATGTG AAATACGAAG GCCGAGCATA CAACACACCT TAACACTCGC CTATTGTATA

BssHII
~~~~~

4901 TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCGC GCAATTAAAC CTCACTAAAG GGAACAAAAG
AGTGTGTCCT TTGTCGATAC TGGTACTAAT GCGGTTCCGC CGTTAATTGG GAGTGATTTC CTTGTTTTC
```

EcoR

4971 CTGG
GACC

FIG..44H

116 / 154

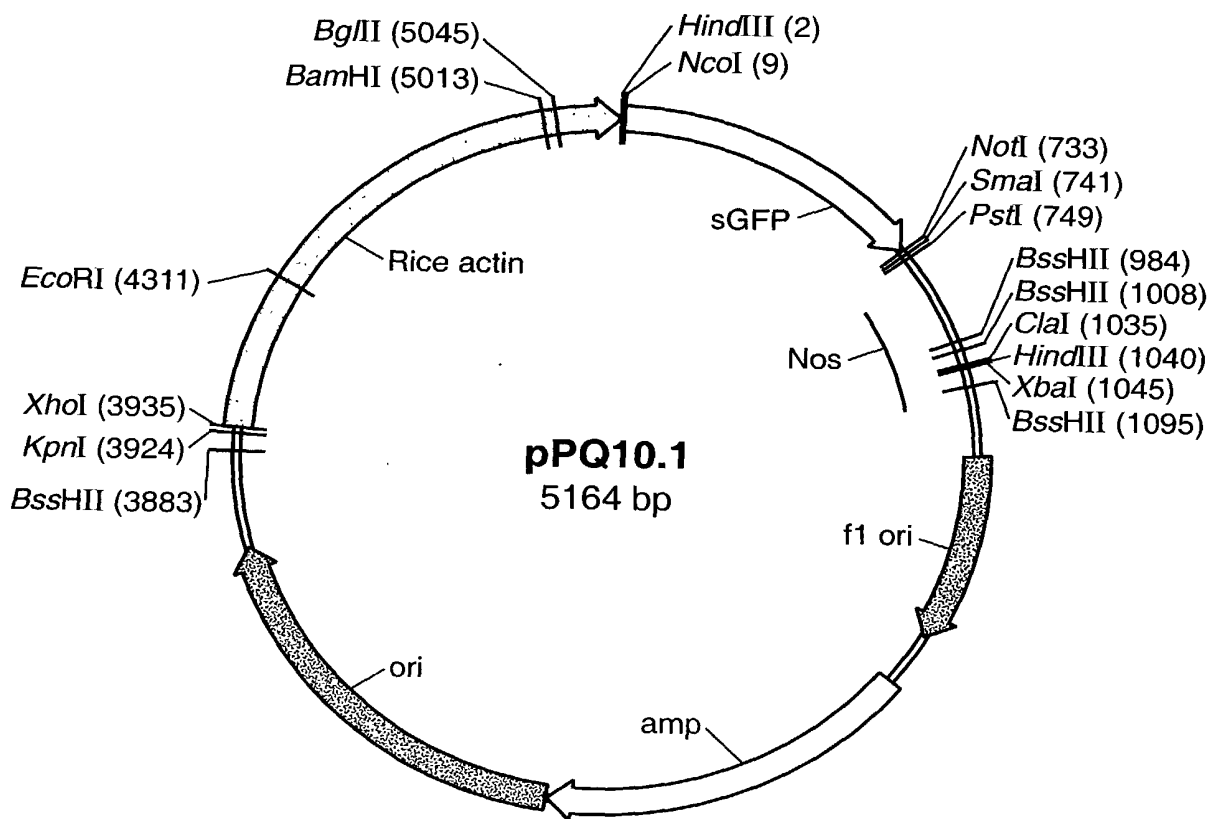


FIG. 45A

117 / 154

HindIII NcoI
~~~~~  
1 AAGCTTACCA TGGTGAGCAA GGGCGAGGAG CTGTTACACG GGTGGTGCC CATCTGGTC GAGCTGGACG  
TTCGAATGGT ACCACTCGTT CCCGCTCCTC GACAAGTGGC CCCACCACGG GTAGGACCAG CTCGACCTGC  
71 GCGACGTGAA CGGCCACAAG TTCAGCGTGT CCGCGGAGGG CGAGGGCGAT GCCACCTACG GCAAGCTGAC  
CGCTGCACTT GCCGGTGTTT AAGTCGCACA GGCCGCTCCC GCTCCCGCTA CCGTGGATGC CGTTCGACTG  
141 CCTGAAGTTC ATCTGCACCA CCGGCAAGCT GCCCGTGCCC TGGCCCAACC TCGTGACCAC CTTACACCTAC  
GGACTTCAAG TAGACGTGGT GGCCGTTTCA CCGGCACGGG ACCGGGTGGG AGCACTGGTG GAAGTGGATG  
211 GCGGTGCAGT GCTTCAGCCG CTACCCCGAC CACATGAAGC AGCACGACTT CTTCAAGTCC GCCATGCCCG  
CCGCACGTCA CGAAGTCGGC GATGGGGCTG GTGTACTTCG TCGTGCTGAA GAAGTTCAGG CCGTACGGGC  
281 AAGGTACGT CCAGGAGCGC ACCATCTTCT TCAAGGACGA CGGCAACTAC AAGACCCGCG CCGAGGTGAA  
TTCCGATGCA GGTCTCTCGG TGGTAGAAGA AGTTCTTGCT GCCGTTGATG TTCTGGGCGC GGCTCCACTT  
351 GTTCGAGGGC GACACCTTGG TGAACCGCAT CGAGCTGAAG GGCATCGACT TCAAGGAGGA CGGCAACATC  
CAAGCTCCC GCTGTGGGACC ACTTGGCGTA GCTCGACTTC CCGTAGCTGA AGTTCTCTCT CCGGTTGTAG  
421 CTGGGGCACA AGCTGGAGTA CAACTACAAC AGCCACAACG TCTATATCAT GGCCGACAAG CAGAAGAACG  
GACCCCGTGT TCGACCTCAT GTTGATGTTG TCGGTGTTGC AGATATAGTA CCGGCTGTTT GTCTTCTTGC  
491 GCATCAAGGT GAACCTCAAG ATCCGCCACA ACATCGAGGA CGGCAGCGTG CAGCTCGCCG ACCACTACCA  
CGTAGTTCCA CTTGAAGTTC TAGGCGGTGT TGTAGCTCCT GCCGTCGCAC GTCGAGCGGC TGGTGATGGT  
561 GCAGAACACC CCCATCGGGC ACGGCCCCGT GCTGCTGCC GACAACCACT ACCTGAGCAC CCAGTCCGCC  
CGTCTTGTGG GGTAGCCGC TGCCGGGGCA CGACGACGGG CTGTTGGTGA TGGACTCGTG GGTACGGCGG  
631 CTGAGCAAAG ACCCCAACGA GAAGCGCGAT CACATGGTCC TGCTGGAGTT CGTGACCGCC GCCGGGATCA  
GACTCGTTTC TGGGGTTGCT CTTCCGCGTA GTGTACCAGG ACGACCTCAA GCACCTGGCG CGGCCCTAGT

FIG.-45B

118 / 154

```

SmaI
~~~~~
NotI PstI
~~~~~
701 CTCACGGCAT GGACGAGCTG TACAAGTAA GCGCCGCCCC GGGCTGCAGG GAAACCACTG AAGGATGAGC
    GAGTCCCGTA CCTGCTCGAC ATGTTCAATT CGCCGGCGGG CCGACGTC CTTTGGTGAC TTCTTACTCG
771 TGTAAAGAAG CAGATCGTTC AACATTTGG CAATAAAGTT TCTTAAGATT GAATCCTGTT GCCGGTCTTG
    ACATTTCTTC GTCTAGCAAG TTTGTAAACC GTTATTTCAA AGAATTCTAA CTTAGGACAA CGGCCAGAAC
841 CGATGATTAT CATATAATTT CTGTTGAATT ACGTTAAGCA TGTAAATAAT AACATGTAAT GCATGACGTT
    GCTACTAATA GTATATTAA GACAACCTAA TGCAATTCGT ACATTATTAA TTGTACATTA CGTACTGCAA
911 ATTTATGAGA TGGGTTTTTA TGATTAGAGT CCCGCAATTA TACATTTAAT ACGCGATAGA AAACAAATA
    TAAATACTCT ACCCAAAAAT ACTAATCTCA GGGCGTTAAT ATGTAAATTA TCGGCTATCT TTTGTTTAT
XbaI
~~~~~
ClaI HindIII
~~~~~
981 TAGCGCGCAA ACTAGGATAA ATTATCGCGC GCGGTGTCAT CTATGTTACT AGATCGATAA GCTTCTAGAG
    ATCGCGCGTT TGATCCTATT TAATAGCGCG CGCCACAGTA GATACAAATGA TCTAGCTATT CGAAGATCTC
BssHII
~~~~~
1051 CGGCCGGTGG AGCTCCAATT CGCCCTATAG TGAGTCGTAT TACGCGCGCT CACTGGCCGT CGTTTACAA
 GCCGGCCACC TCGAGGTTAA GCGGGATATC ACTCAGCATA ATGCGCGCGA GTGACCGGCA GCAAAATGTT
1121 CGTCGTGACT GGGAAAACCC TGGCGTTACC CAACTTAATC GCCTTGCAGC ACATCCCCCT TTCGCCAGCT
 GCAGCACTGA CCCTTTTGG ACCGCAATGG GTTGAATTAG CGGAACGTCG TGTAGGGGGA AAGCGGTGCA
1191 GGC GTAATAG CGAAGAGGCC CGCACCAGATC GCCCTTCCCA ACAGTTGCGC AGCCTGAATG GCGAATGGGA
 CCGCATATC GCTTCTCCGG GCGTGGCTAG CGGGAAGGCT TGTC AACGCG TCGGACTTAC CGCTTACCCCT

```

FIG._45C

1261 CGCGCCCTGT AGCGGCGCAT TAAGCGCGGC GGGTGTGGTG GTTACGCGCA GCGTGACCGC TACACTTGCC  
GCGCGGGACA TCGCCGCGTA ATTGCGCGCG CCCACACCAC CAATGCGCGT CGCACTGGCG ATGTGAACGG

1331 AGCGCCCTAG CGCCCGCTCC TTTCGCTTTC TTCCCTTCCT TTCTCGCCAC GTTCGCCGGC TTTCCCGCTC  
TCGCGGGATC GCGGGCGAGG AAAGCGAAAG AAGGAAGGA AAGAGCGGTG CAAGCGGCCG AAAGGGGCAG

1401 AAGCTCTAAA TCGGGGGCTC CCTTTAGGGT TCCGATTAG TGCTTTACGG CACCTCGACC CCAAAAAACT  
TTCGAGATT T AGCCCCCGAG GGAATCCCA AGGCTAAATC ACGAAATGCC GTGGAGCTGG GGTTTTGA

1471 TGATTAGGGT GATGTTTAC GTAGTGGGC ATCGCCCTGA TAGACGGT TTCGCCCTTT GACGTTGGAG  
ACTAATCCCA CTACCAAGTG CATCACCCGG TAGCGGGACT ATCTGCCAAA AAGCGGGA AAGCTCAACCTC

1541 TCCACGTTCT TTAATAGTGG ACTCTTGTT CAAACTGGAA CAACACTCAA CCTATCTCG GTCTATTCTT  
AGGTGCAAGA AATTATCACC TGAGAACAAG GTTTGACCTT GTTGAGGT GGGATAGAGC CAGATAAGAA

1611 TTGATTATA AGGGATTTG CCGATTTCGG CCTATTGGT AAAAATGAG CTGATTAAAC AAAAATTAA  
AACTAAATAT TCCCTAAAAC GGCTAAAGCC GGATAACCAA TTTTITACTC GACTAAATG TTTTITAAAT

1681 CGCGAATTT AACAAAAATAT TAACGCTTAC AATTAGGT GCACITTTTCG GGGAAATGTG CGCGGAACCC  
GCGCTTAAAA TTGTTTATA ATTGCGAATG TTAATCCAC CGTGAAAAGC CCTTTTACAC GCGCCTTGGG

1751 CTATTGTGT ATTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCCT GATAAATGCT  
GATAAACAAA TAAAAAGATT TATGTAAGT TATACATAGG CGAGTACTCT GTTATTGGGA CTATTACGA

1821 TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCTG CCTTATTCC CTTTITGCG  
AGTTATTATA ACTTTTCTCT TCTCATACTC ATAAGTTGTA AAGGCACAGC GGAATAAGG GAAAAACCG

1891 GCATTTGCG TTCTGTGTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA AGATGCTGAA GATCAGTTGG  
CGTAAAACGG AAGGACAAA ACGAGTGGGT CTTTGGGACC ACTTTCATTT TCTACGACTT CTAGTCAACC

1961 GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTC GCCCCGAAGA  
CACGTGCTCA CCCAATGTAG CTTGACCTAG AGTTGTGCGC ATTCTAGGAA CTCTCAAAAG CCGGGCTTCT

2031 ACGTTTCCA ATGATGAGCA CTTTAAAGT TCTGCTATGT GCGCGGTAT TATCCCGTAT TGACGCCGGG  
TGCAAAAGGT TACTACTCGT GAAATTTCA AGACGATACA CCGCGCCATA ATAGGCAATA ACTGCGGCC

FIG. 45D

2101 CAAGAGCAAC TCGGTCGCCG CATACTAT TCTCAGAATG ACTTGTTGA GTACTACCA GTACACAGAAA  
GTTCTCGTTG AGCCAGCGGC GTATGTGATA AGAGTCTTAC TGAACCAACT CATGAGTGGT CAGTGTCTTT

2171 AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC  
TCGTAGAATG CCTACCGTAC TGTCACTCTC TTAATACGTC ACGACGGTAT TGGTACTCAC TATTGTGACG

2241 GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA CATGGGGGAT  
CCGGTTGAAT GAAGACTGTT GCTAGCCTCC TGGCTTCCTC GATTGGCGAA AAAACGTGTT GTACCCCTTA

2311 CATGTAATC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA  
GTACATTGAG CGGAACCTAGC AACCCCTTGGC CTCGACTTAC TTCGGTATGG TTTGCTGCTC GCACTGTGGT

2381 CGATGCCCTGT AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA CTACTTACTC TAGCTTCCCG  
GCTACGGACA TCGTTACCGT TGTGCAACG CGTTTGATAA TTGACCGCTT GATGAATGAG ATCGAAGGGC

2451 GCAACAATTA ATAGACTGGA TGGAGCGGA TAAAGTTGCA GGACCACTTC TGGCTCGGC CCTTCCGGCT  
CGTTGTTAAT TATCTGACCT ACCTCCGCTT ATTTCAACGT CCTGGTGAAG ACGCGAGCCG GGAAGGCCGA

2521 GGCTGGTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCG TATCATTTGCA GCACTGGGGC  
CCGACCAAAT AACGACTATT TAGACCTCGG CCACTCGCAC CCAGAGCGCC ATAGTAACGT CGTGACCCCG

2591 CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA  
GTCTACCAAT CGGGAGGGCA TAGCATCAAT AGATGTGCTG CCCCCTCAGTC CGTTGATACC TACTTGCTTT

2661 TAGACAGATC GCTGAGATAG GTGCCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT TTACTCATAT  
ATCTGTCTAG CGACTCTATC CACGGAGTGA CTAATTCGTA ACCATTGACA GTCTGGTTCA AATGAGTATA

2731 ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT GAAGATCCTT TTTGATAATC  
TATGAAATCT AACTAAATTT TGAAGTAAAA ATTAAATTTT CCTAGATCCA CTTCTAGGAA AAACATATTAG

2801 TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG  
AGTACTGGTT TTAGGGAAAT GCACTCAAAA GCAAGGTGAC TCGCAGTCTG GGGCATCTTT TCTAGTTTCC

2871 ATCTTCTTGA GATCCTTTTT TTCTGGCGGT AATCTGCTGC TTGCAAAACAA AAAAACACC GCTACCAGCG  
TAGAAGAACT CTAGGAAAAA AAGACGCGCA TTAGACGAGC AACGTTTGT TTTTGTGGTGG CGATGGTCCG

FIG._45E



121 / 154

2941 GTGGTTTGTG TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA  
CACCAAAACA ACGGCCTAGT TCTCGATGGT TGAGAAAAAG GCTTCCATTG ACCGAAGTCG TCTCGCGTCT

3011 TACCAAATAC TGTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC  
ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC TTGAGACATC GTGGCGGATG

3081 ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGCCGATA AGTCGTGTCT TACCGGGTTG  
TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG TCACCGCTAT TCAGCACAGA ATGGCCCAAC

3151 GACTCAAGAC GATAGTTACC GGATAAGCG CAGCGGTCCG GCTGAACGGG GGGTTCGTGC ACACAGCCCA  
CTGAGTTCTG CTATCAATGG CCTATTCCGC GTCGCCAGCC CGACTTGCCC CCCAAGCACG TGTGTCGGGT

3221 GCTTGGAGCG AACGACCTAC ACCGAACCTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG CCACGCTTCC  
CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGAT ACTCTTTCGC GGTGCGAAGG

3291 CGAAGGGAGA AAGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT  
GCTTCCCTCT TTCCGCCTGT CCATAGGCCA TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG CTCCCTCGAA

3361 CCAGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT  
GGTCCCCCTT TCGGACCCT AGAAATATCA GGACAGCCCA AAGCGGTGGA GACTGAATC GCAGCTAAAA

3431 TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTTC GGTTCCTGGC  
ACACTACGAG CAGTCCCCCC GCCTCGGATA CCTTTTTCG GTCGTTGCGC CGGAAAAATG CCAAGGACCG

3501 CTTTTGCTGG CTTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA CCGTATTACC  
GAAAAACGAC GAAAAACGAG TGTACAAGAA AGGACGCAAT AGGGGACTAA GACACCTATT GGCATAATGG

3571 GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG  
CGGAAACTCA CTCGACTATG GCGAGCGCGG TCGGCTTGCT GGCTCGCGTC GCTCAGTCAC TCGCTCCTTC

3641 CGGAAGAGCG CCCAATACG AACC CGCCTC TCCCGCGCG TTGGCCGATT CATTAATGCA GCTGGCACGA  
GCCTTCTCGC GGGTTATGCG TTGCGCGGAG AGGGCGCGC AACCGGCTAA GTAATTACGT CGACCGTGCT

3711 CAGGTTTCCC GACTGGAAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA GTTAGCTCAC TCATTAGGCA  
GTCCAAAGGG CTGACCCTTC GCCCGTCACT CGCGTTGCGT TAATTACACT CAATCGAGTG AGTAATCCGT

FIG._45F

122 / 154

```
3781 CCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATGTTGT GTGGAATTGT GAGCGGATAA CAATTTACACA
GGGTCCGAA ATGTGAAATA CGAAGGCCGA GCATACAACA CACCTTAACA CTCGCCATT CTGAAAGTGT

3851 CAGGAACAG CTATGACCAT GATTACGCCA AGCGCGCAAT TAACCCCTCAC TAAAGGGAAC AAAAGCTGGG
GTCCTTTGTC GATACTGGTA CTAATGCGGT TCGCGCGTTA ATTGGGAGTG ATTTCCCTTG TTTTCGACCC

 BssHII
      ~~~~~
      KpnI
      ~~~

3921 TACCGGGCCC CCCCTCGAGG TCATTCAAT GCTTGAGAAG AGAGTCGGGA TAGTCCAAAA TAAACAAAAG
ATGGCCCGGG GGGGAGCTCC AGTAAGTATA CGAACTCTTC TCTCAGCCCT ATCAGGTTTT ATTTTGTTC

3991 GTAAGATTAC CTGGTCAAAA GTGAAAACAT CAGTTAAAAG GTGGTATAAG TAAAATATCG GTAATAAAAG
CATTCTAATG GACCAGTTTT CACTTTTGTA GTCAAATTTT CACCATATTC ATTTATATAG CATTATTTTC

4061 GTGGCCCAA GTGAAATTTA CTCTTTTCTA CTATTATAAA AATTGAGGAT GTTTTGTCTG TACTTTGATA
CACCGGGTTT CACTTTAAAT GAGAAAAGAT GATAATATTT TTAACCTCTA CAAAACAGCC ATGAAACTAT

4131 CGTCATTTT GTATGAATG GTTTTAAAGT TTATTCGCGA TTTGGAAATG CATATCTGTA TTTGAGTCGG
GCAGTAAAA CATACTTAAC CAAAAATTCA AATAAGCGCT AAACCTTTAC GTATAGACAT AAACCTCAGCC

4201 TTTTAAAGT CGTTGCTTTT GTAAATACAG AGGGATTGT ATAAGAAATA TCTTTAAAAA ACCCATATGC
AAAAATTCAA GCAACGAAAA CATTATATGC TCCCTAAACA TATTCTTTAT AGAAATTTTT TGGGTATACG

 EcoRI
      ~~~~~

4271 TAATTGACA TAATTTTGA GAAAAATATA TATTCAGGCG AATCCACAA TGAACAATAA TAAGATTAAA
ATTAAACTGT ATTAAAAACT CTTTTATAT ATAAAGTCCG TTAAGGTGTT ACTGTTATT ATCTAATTT

4341 ATAGCTTGC CCCGTTGCAG CGATGGGTAT TTTTCTAGT AAAATAAAAG ATAAACTTAG ACTCAAAACA
TATCGAACGG GGGCAACGTC GCTACCCATA AAAAAGATCA TTTTATTTTC TATTGAATC TGAGTTTTGT

4411 TTTACAAAA CAACCCCTAA AGTCCTAAAG CCCAAAGTGC TATGCACGAT CCATAGCAAG CCCAGCCCCA
AAATGTTTT GTTGGGGATT TCAGGATTTT GGGTTTCAG ATACGTGCTA GGTATCGTTC GGTCTGGGTT
```

FIG..45G

123 / 154

```
4481 CCCAACCCAA CCCAACCCAC CCCAGTGCAG CCAACTGGCA AATAGTCTCC ACCCCCGGCA CTATCACCGT
GGGTTGGGTT GGGTTGGGTG GGGTCACGTC GGTTGACCGT TTATCAGAGG TGGGGGCCGT GATAGTGGCA

4551 GAGTTGTCCG CACCACCGCA CGTCTCGCAG CCAAAAAAGG AAAAAAAGG AAAAAAAGG
CTCAACAGGC GTGGTGGCGT GCAGAGCGTC GGTTTTCTTT TTTTCTTTCT TTTTCTTTT

4621 CAGCAGGTGG GTCCGGGTG GTGGGGCCCG AAAAGCGAGG AGGATCGCGA GCAGCGACGA GGCCCGGGCC
GTCGTCCACC CAGGCCCAGC ACCCCCGGCC TTTTTCGCTCC TCCTAGCGCT CGTCGCTGCT CCGGGCCCGG

4691 TCCCTCCGCT TCCAAAGAAA CGCCCCCCAT CGCCACTATA TACATACCCC CCCCTCTCCT CCCATCCCCC
AGGGAGGCGA AGGTTCTTT GCGGGGGGTA GCGGTGATAT ATGTATGGG GGGAGAGGA GGGTAGGGG

4761 CAACCTACC ACCACCACCA CCACCACCTC CTCCCCCCTC GGTGCCGGAC GACGAGCTCC TCCCCCCTCC
GTTGGGATGG TGTGGTGGT GGTGGTGGAG GAGGGGGGAG CGACGGCCTG CTGCTCGAG AGGGGGGAGG

4831 CCCTCCGCGG CCGCCGGTAA CCACCCCGCC CCTCTCCTCT TTCTTTCTCC GTTTTCTTTT TCGTCTCGGT
GGGAGGCGGC GCGGGCCATT GGTGGGCGG GGAGAGGAGA AAGAAAGAGG CAAAAAAGG AGCAGAGCCA

4901 CTCGATCTTT GGCCTTGGTA GTTTGGGTGG GCGAGAGCGG CTTCGTCGCC CAGATCGGTG CGCGGGAGGG
GAGCTAGAAA CCGGAACCAT CAAACCCACC CGCTCTCGCC GAAGCAGCGG GTCTAGCCAC GCGCCCTCCC

          BamHI
          ~~~~~

4971 GCGGGATCTC GCGGCTGGCG TCTCCGGGCG TGAGTCGGCC CGGATCCTCG CGGGGAATGG GGCTCTCGGA
CGCCCTAGAG CGCCGACCGC AGAGGCCCGC ACTCAGCCCG GCCTAGGAGC GCCCCTTACC CCGAGAGCCT

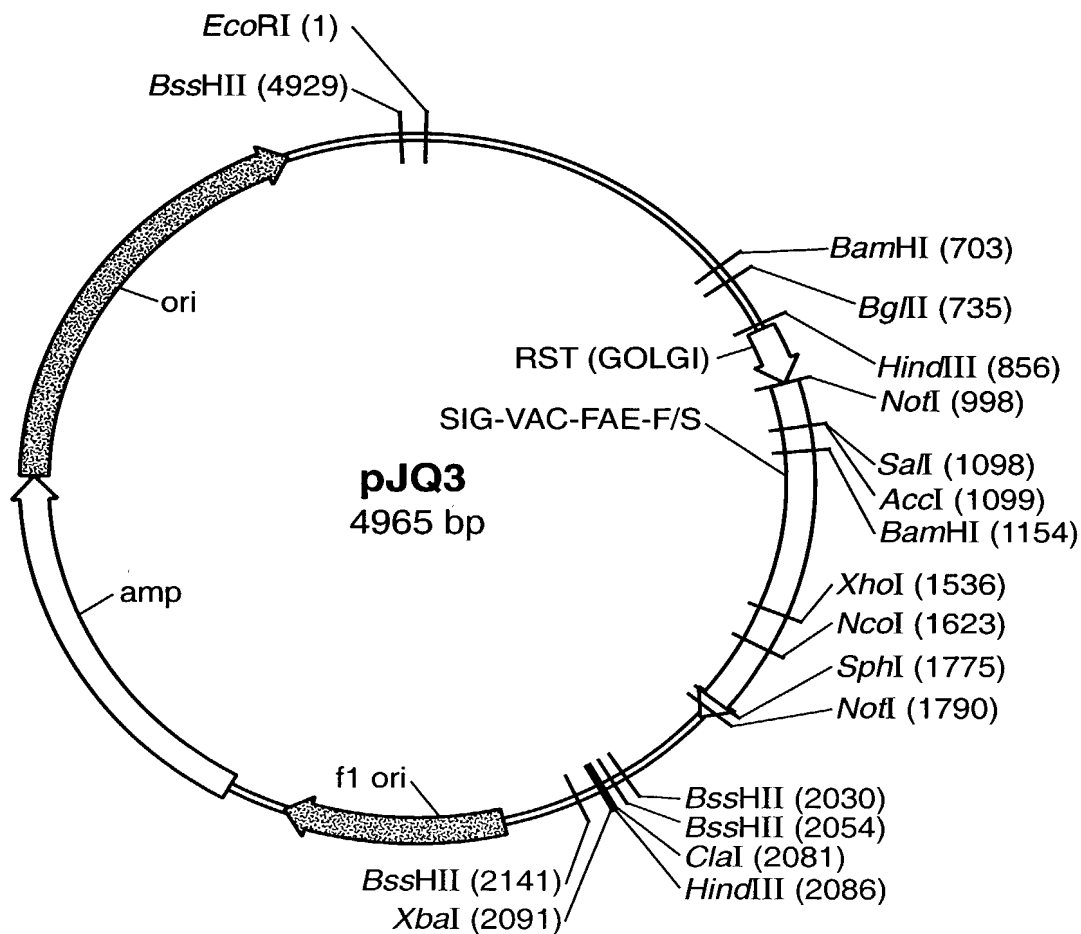
 BglII
          ~~~~~

5041 TGTAGATCTT CTTTCTTTCT TCTTTTGTG GTAGAAATTG AATCCCTCAG CATGTTCAT CGGTAGTTTT
ACATCTAGAA GAAAGAAAGA AGAAAAACAC CATCTTAAAC TTAGGGAGTC GTAACAAGTA GCCATCAAAA

5111 TCTTTTCATG ATTTGTGACA AATGCAGCCT CGTGGGGAGC TTTTGTGTAG GTAG
AGAAAGTAC TAAACACTGT TTACGTGCGA GCACGCCCTG AAAAAACATC CATC
```

FIG._45H

124 / 154



**FIG._46A**

125 / 154

```

ECORI
~~~~~
1 AATTCACAA TGAACAATAA TAAGATTAAA ATAGCTTGCC CCCGTTGCAG CGATGGGTAT TTTTCTAGT
 TTAAGGTGTT ACTTGTTATT ATCTAATTT TATCGAACGG GGGCAACGTC GCTACCCATA AAAAAGATCA

71 AAAATAAAAG ATAAACTTAG ACTCAAAAACA TTTACAAAAA CAACCCCTAA AGTCCCTAAAG CCCAAAGTGC
 TTTTATTTTC TATTGAATC TGAGTTTGT AAATGTTTTT GTTGGGGATT TCAGGATTC GGGTTTCACG

141 TATGCACGAT CCATAGCAAG CCAGCCCCAA CCCAACCCAC CCCAGTGCAG CCAACTGGCA
 ATACGTGCTA GGTATCGTTC GGTTCGGGTT GGTTCGGTG GGTTCACGTC GGTGACCGT

211 AATAGTCTCC ACCCCCGGCA CTATCACCGT GAGTGTCCG CACCACCGCA CGTCTCGCAG CCAAAAAAAA
 TTATCAGAGG TGGGGGCCGT GATAGTGGCA CTCAACAGGC GTGGTGGCGT GCAGAGCGTC GTTTTTTTTT

281 AAAAAGAAAG AAAAAAAGA AAAAGAAAAA CAGCAGGTGG GTCCGGGTGG TGGGGGCCGG AAAAGCGAGG
 TTTTCTTTC TTTTCTTTC TTTTCTTTC GTCGTCCACC CAGCCCCAGC ACCCCCGGCC TTTTCGCTCC

351 AGGATCGCGA GCAGCGACGA GGCCCGGCC TCCCTCCGCT TCCAAAGAAA CGCCCCCCAT CGCCACTATA
 TCCTAGCGCT CGTCGCTGCT CCGGCCCGG AGGAGGCCGA AGTTTCTTT GCGGGGGGTA GCGGTGATAT

421 TACATACCCC CCCCTCTCCT CCCATCCCC CAACCTTACC ACCACCACTC CCACCACTC CTCCCCCTC
 ATGATGGGG GGGGAGAGGA GGTAGGGGG GTTGGGATGG TGGTGGTGGT GGTGGTGGAG GAGGGGGGAG

491 GCTGCCGGAC GACGAGCTCC TCCCCCTCC CCTCCGCGG CCGCCGGTAA CCACCCCGCC CCTCTCTCT
 CGACGGCCTG CTGCTCGAG AGGGGGGAG GGGAGCGGC GCGGCCCAT GTTGGGGCGG GGAGAGGAGA

561 TTCTTTCTCC GTTTTTTTTT TCGTCTCGGT CTCGATCTTT GGCTTGGTA GTTGGGTGG GCGAGAGCGG
 AAGAAAGAGG CAAAAAAA AGCAGAGCCA GAGCTAGAAA CCGGAACCAT CAAACCCACC CGTCTCTGCC

631 CTTCTGTCGC CAGATCGGTG CCGGGGAGG GCGGATCTC GCGCTGGCG TCTCCGGCG TGAGTCGGCC
 GAAGCAGCGG GTCTAGCCAC GCGCCCTCC CGCCCTAGAG CCGGACCGC AGAGGCCCGC ACTCAGCCGG

BamHI
~~~~~
701 CGGATCCTCG CGGGGAATGG GGCTCTCGGA TGATGATCTT CTTTCTTCT TCTTTTGTG GTAGAATTG
   GCCTAGGAGC GCCCTTACC CCGAGAGCCT ACATCTAGAA GAAAGAAAGA AGAAAAACAC CATCTTAAAC

BglII
~~~~~

```

FIG..46B

126 / 154

```
771 AATCCCTCAG CATTGTTTCAT CGGTAGTTTTT TCTTTTCATG ATTTGTGACA AATGCAGCCT CGTGGGGAGC
TTAGGGAGTC GTAACAAGTA GCCATCAAAA AGAAAAGTAC TAAACACTGT TTACGTCGGA GCACGCCCTCG

HindIII
~~~~~
841 TTTTTTTGTAG GTAGAAGCTT ACCATGATCC ACACCAACCT CAAAAGAAG TTCTCCCTCT TCATCCTCGT
AAAAACATC CATCTTCGAA TGGTACTAGG TGTGGTTGGA GTTTTCTTC AAGAGGGAGA AGTAGGAGCA

911 CTTCCTCCTC TTCGCCGTGA TCTGCGTGTG GAAAGAAGGC TCCGACTACG AGGCCCTCAC CCTCCAAGCC
GAAGGAGGAG AAGCGGCACT AGACGCACAC CTTCCTCCCG AGGCTGATGC TCCGGGAGTG GGAGGTTCCG

NotI
~~~~~
981 AAGGAGTTCC AAATGGCGGC CGCCTCCACG CAGGGCATCT CCGAAGACCT CTACAGCCGT TTAGTCGAAG
TTCTCTCAAGG TTTACCGCCG GCGGAGGTGC GTCCCCGTAG GGCTTCTGGA GATGTCGGCA AATCAGCTTT

SalI
~~~~~
AccI
~~~~~
1051 TGGCCACTAT CTCCCAAGCT GCCTACGCCG ACCTGTGCAA CATTCCGTCG ACTATTATCA AGGAGAGAA
ACCGGTGATA GAGGGTTCGA CGGATGCGGC TGGACACGTT GTAAGGCAGC TGATAATAGT TCCCTCTCTT

BamHI
~~~~~
1121 AATTACAAT TCTCAAACATG ACATTAAACGG ATGGATCCTC CGCGACGACA GCAGCAAAGA AATAATCACC
TTAAATGTTA AGAGTTTGAC TGTAATTGCC TACCTAGGAG GCGCTGCTGT CGTCGTTTCT TTATTAGTGG

1191 GTCTTCCGTG GCACTGGTAG TGATACGAAT CTACAACTCG ATACTAACTA CACCCCTCACG CCTTTCGACA
CAGAAGGCAC CGTGACCATC ACTATGCTTA GATGTTGAGC TATGATTGAT GTGGGAGTGC GGAAAGCTGT

1261 CCCTACCACA ATGCAACGGT TGTGAAGTAC ACGGTGGATA TTATATTGGA TGGGTCTCCG TCCAGGACCA
GGGATGGTGT TACGTTGCCA ACACCTTCATG TGCCACCTAT AATATAACCT ACCCAGAGGC AGGTCCTGGT
```

FIG._46C

127 / 154

```
1331 AGTCGAGTCG CTTGTCAAAC AGCAGGTTAG CCAGTATCCG GACTACGCGC TGACCGTGAC CGGCCACKCC
    TCAGCTCAGC GAACAGTTTG TCGTCCAATC GGTCATAGGC CTGATGCGCG ACTGGCACTG GCCGGTGMGG

1401 CTCGGCGCCT CCCTGGCGGC ACTCACTGCC GCCCAGCTGT CTGCGACATA CGACAACATC CGCCTGTACA
    GAGCCGCGGA GGGACCGCCG TGAGTGACGG CGGGTCGACA GACGCTGTAT GCTGTTGTAG GCGGACATGT

                                XhoI
                                ~~~~~
1471 CCTTCGGCGA ACCGCGCAGC GGCAATCAGG CCTTCGCGTC GTACATGAAC GATGCCCTTC AAGCCTCGAG
 GGAAGCCGCT TGGCGCGTCG CCGTTAGTCC GGAAGCGCAG CATGTACTTG CTACGGAAGG TTCGGAGCTC

1541 CCCAGATACG ACGCAGTATT TCCGGGTCAC TCATGCCAAC GACGGCATCC CAAACCTGCC CCCGGTGGAG
 GGGTCTATGC TGCGTCAATA AGGCCCAGTG AGTACGGTTG CTGCCGTAGG GTTTGGACGG GGGCCACCCTC

 NcoI
                                ~~~~~
1611 CAGGGGTACG CCCATGGCGG TGTAGAGTAC TGGAGCGTTG ATCCTTACAG CGCCCAGAAC ACATTTGTCT
    GTCCCCATGC GGTACCGCC ACATCTCATG ACCTCGCAAC TAGGAATGTC GCGGTCCTTG TGTAACACAGA

1681 GCACTGGGGA TGAAGTGCAG TGCTGTGAGG CCCAGGGCGG ACAGGGTGTG AATAATGCGC ACACGACTTA
    CGTGACCCCT ACTTCACGTC ACGACACTCC GGTCCCCGCC TGTCCCACAC TTATTACGCG TGTGCTGAAT

                                SphI
                                ~~~~~
 NotI
                                ~~~~~
1751 TTTTGGGATG ACGAGCGGCG CATGCACCTG GCCGGTCGCG GCCCGGAAA CCACTGAAGG ATGAGCTGTA
    AAAACCTAC TGCTCGCCCG GTACGTGGAC CGGCCAGCGC CGGCCCTTT GGTGACTTCC TACTCGACAT

1821 AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT AAAGTTTCTT AAGATTGAAT CCTGTTGCCG GTCTTGCGAT
    TTCTTCGTCT AGCAAGTTTG TAAACCGTTA TTTCAAAGAA TTCTAACTTA GGACAACGGC CAGAACGCTA

1891 GATTATCATA TAATTCTGT TGAATTACGT TAAGCATGTA ATAATTACA TGTAAATGCAT GACGTTATTT
    CTAATAGTAT ATTAAGACA ACTTAATGCA ATTCTGTACAT TATTAATTGT ACATTACGTA CTGCAATAAA
```

FIG._46D

128 / 154

```

BssHII
1961  ATGAGATGGG TTTTATTGAT TAGAGTCCCG CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC
      TACTCTACCC AAAAATACTA ATCTCAGGCG GTTAATATGT AAATTATGCG CTATCTTTTG TTTTATATCG
      ~~~~~
 XbaI
      ~~~~~
      ClaI HindIII
2031  GCGCAAACTA GGATAAATTA TCGCGCGCGG TGTCATCTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC
      CGCGTTTGAT CCTATTTAAT AGCGCGCGCC ACAGTAGATA CAATGATCTA GCTATTCGAA GATCTCGGCC
      ~~~~~
 BssHII
      ~~~~~
2101  CGGTGGAGCT CCAATTGCGC CTATAGTGAG TCGTATTACG CGCGCTCACT GGCCGTCGTT TTACAACGTC
      GCCACCTCGA GGTAAAGCGG GATATCACTC AGCATAATGC GCGCGAGTGA CCGGCAGCAA AATGTTGCAG
      ~~~~~
2171 GTGACTGGGA AAACCCCTGC GTTACCCAACT TTAATCGCCT TGCAGCACAT CCCCCTTTTCG CCAGCTGGCG
 CACTGACCCCT TTTGGGACCG CAATGGGTTG AATTAGCGGA ACCTCGTGTA GGGGAAAGC GGTCGACCCG
      ~~~~~
2241  TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG
      ATTATCGCTT CTCCGGGCGT GGCTAGCGGG AAGGTTGTC AACGCGTCGG ACTTACCGCT TACCCTGCGC
      ~~~~~
2311 CCCTGTAGCG GCGCATTAAG CCGGCGGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG
 GGGACATCGC CGCGTAATTC GCGCCGCCCA CACCACCAAT GCGCGTCGCA CTGGCGATGT GAACGGTCGC
      ~~~~~
2381  CCCTAGCGCC CGCTCCTTTC GCTTCTTCC CTTCTTTCT CGCCACGTTT CCGGGCTTTC CCCGTCAAAG
      GGGATCGCGG GCGAGGAAAG CGAAAGAAAG GAAGGAAAGA GCGGTGCAAG CGGCCGAAAG GGGCAGTTGC
      ~~~~~
2451 TCTAAATCGG GGGCTCCCTT TAGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT
 AGATTTAGCC CCCGAGGGA ATCCCAAGGC TAAATCACGA AATGCCGTGG AGCTGGGGTT TTTTGAACTA
      ~~~~~
2521  TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTTCG CCCTTTGACG TTGGAGTCCA
      ATCCCACTAC CAAGTGCATC ACCCGGTAGC GGGACTATCT GCCAAAAAGC GGGAAACTGC AACCTCAGGT

```

FIG..46E



129 / 154

2591 CGTCTCTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCCCT ATCTCGGTCT ATTCCTTTGA  
GCAAGAAATT ATCACCTGAG AACAAGGTTT GACCTTGTTG TGAGTTGGGA TAGAGCCAGA TAAGAAAACT

2661 TTTATAAGGG ATTTTGCCGA TTTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTTTAACAAA ATTTAACGCG  
AAATATTCCC TAAAACGGCT AAAGCCGGAT AACCAATTTT TTAAGCTGACT AAATGTTT TAAATGCGC

2731 AATTTTAAAC AAATATTAAAC GCTTACAATT TAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT  
TTAAAAATTGT TTTATAATTG CGAATGTTAA ATCCACCGTG AAAAGCCCCCT TTACACGCGC CTTGGGGATA

2801 TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCTTGATA AATGCTTCAA  
AACAAATAAA AAGATTATG TAAGTTTATA CATAGCGGAG TACTCTGTTA TTGGGACTAT TTACGAAAGTT

2871 TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTC CCATGCGCCCT TATTCCTCTT TTTGCGGCAT  
ATTATAACTT TTTCCCTTCTC ATACTCATAA GTTGTAAGG CACAGCGGGA ATAAGGGAAA AAACGCCGTA

2941 TTTGCCCTCC TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC  
AAACGGAAGG ACAAACGA GTGGGTCTTT GCGACCACTT TCATTTCTA CGACTTCTAG TCAACCCACG

3011 ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT  
TGCTCACCCA ATGTAGCTTG ACCTAGAGTT GTCGCCATTC TAGGAACCTCT CAAAAGCGGG GCTTCTTGCA

3081 TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTATTGAC GCCGGGCAAG  
AAAGGTTACT ACTCGTGAAA ATTTCAAGAC GATACACCGC GCCATAATAG GGCATAACTG CGGCCCGTTC

3151 AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTGAGTAC TCACCAGTCA CAGAAAAGCA  
TCGTTGAGCC AGCGGCGTAT GTGATAAGAG TCTTACTGAA CCAACTCATG AGTGGTCAGT GTCTTTTCGT

3221 TCTTACGGAT GGCATGACAG TAAGAGAAAT ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC  
AGAAATGCCA CCGTACTGTC ATTCCTTTAA TACGTACCGA CCGTATTGGT ACTCACTATT GTGACGCCCG

3291 AACTTACTTC TGACAAACGAT CGGAGGACCG AAGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG  
TTGAATGAAG ACTGTTGCTA GCCTCCTGGC TTCCCTCGATT GGCGAAAAAA CGTGTGTGAC CCCCTAGTAC

3361 TAACCTGCCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAC GACGAGCGTG ACACCACGAT  
ATTGAGCGGA ACTAGCAACC CTTGGCCCTCG ACTTACTTCG GTATGGTTTG CTGCTCGCAC TGTGTGCTA

FIG._46F

130 / 154

3431 GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT GCGAACTAC TTACTCTAGC TTCCCGGCAA  
CGGACATCGT TACCGTTGTT GCAACGCGTT TGATAATTGA CCGCTTGATG AATGAGATCG AAGGGCCGTT

3501 CAATTAATAG ACTGGATGGA GCGGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT  
GTTAATTATC TGACCTACCT CCGCCTATT TCAACGTCCGT GTGAAGACGC GAGCCGGGAA GGCCGACCGA

3571 GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA  
CCAAATAACG ACTATTTAGA CCTCGGCCAC TCGCACCCAG AGCGCCATAG TAACGTCGTG ACCCCGGTCT

3641 TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA  
ACCATTCGGG AGGCATAGC ATCAATAGAT GTGCTGCCCC TCAGTCCGTT GATACCTACT TGCTTTATCT

3711 CAGATCGCTG AGATAGGTGC CTCACGTGATT AAGCATTGGT AACTGTGAGA CCAAGTTTAC TCATATATAC  
GTCTAGCGAC TCTATCCACG GAGTGACTAA TTCGTAACCA TTGACAGTCT GGTCAAATG AGTATATATG

3781 TTTAGATTGA TTTAAAACCTT CATTTTTAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTGT ATAATCTCAT  
AAATCTAACT AAATTTTGAA GTAAAAATTA AATTTTCCCTA GATCCACTTC TAGGAAAAAC TATTAGAGTA

3851 GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAAGAT CAAAGGATCT  
CTGGTTTITAG GGAATTGCAC TCAAAAAGCAA GGTGACTCGC AGTCTGGGGC ATCTTTTCTA GTTTCCTAGA

3921 TCTTGAGATC CTTTTTITCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG  
AGAACTCTAG GAAAAAAGA CGCGCATTAG ACGACGAACG TTTGTTTTIT TGGTGGCGAT GGTGCGCCACC

3991 TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTTCGAA GGTAACCTGGC TTCAGCAGAG CGCAGATACC  
AAACAAACGG CCTAGTTCTC GATGGTTGAG AAAAAGGCTT CCATTGACCG AAGTCGTCTC GCGTCTATGG

4061 AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAAGT CTGTAGCACC GCCTACATAC  
TTTATGACAG GAAGATCACA TCGGCATCAA TCCGGTGGTG AAGTTCTTGA GACATCGTGG CGGATGTATG

4131 CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT  
GAGCGAGACG ATTAGGACAA TGGTCACCGA CGACGGTCCAC CGCTATTTCAG CACAGAATGG CCAACCTGA

4201 CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTGCGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT  
GTTCTGCTAT CAATGGCCCTA TTCCGCGTCTG CCAGCCCCGAC TTGCCCCCCA AGCACGTGTG TCGGGTCGAA

FIG._46G

131 / 154

```
4271 GGAGCGAACG ACCTACACCG AACTAGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCAGAA
    CCTCGCTTGC TGGATGTGGC TTGACTCTAT GGATGTCGCA CTCGATACTC TTTCGCGGTG CGAAGGGCTT

4341 GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG
    CCTCTTTCC GCGTGTCAT AGGCCATTCTG CCGTCCCAGC CTTGTCTCTCT CGCGTGCTCC CTCGAAGGTC

4411 GGGGAAACGC CTGGTATCTT TATAGTCCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTGTG
    CCCCCTTTCG GACCATAGAA ATATCAGGAC AGCCCAAAGC GGTGGAGACT GAACTCGCAG CTAATAACAC

4481 ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT
    TACGAGCAGT CCCCCCGCCT CGGATACCTT TTTGCGGTCG TTGCGCCGGA AAAATGCCAA GGACCGGAAA

4551 TGCTGGCCTT TTGCTCACAT GTTCTTTCTT CCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT
    ACGACCGGAA AACGAGTGTA CAAGAAAGGA CGCAATAGGG GACTAAGACA CCTATTGGCA TAATGGCGGA

4621 TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCAG GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA
    AACTCACTCG ACTATGGCGA GCGGCGTCGG CTTGCTGGCT CGCGTCGCTC AGTCACCTCG TCCTTCGCCT

4691 AGAGCGCCCA ATACGCAAAC CGCCTCTCCC CGCGCGTTGG CCGATTCAAT AATGCAGCTG GCACGACAGG
    TCCTCGCGGT TATGCGTTTG GCGGAGAGGG GCGCGCAACC GGCTAAGTAA TTACGTCGAC CGTGCTGTCC

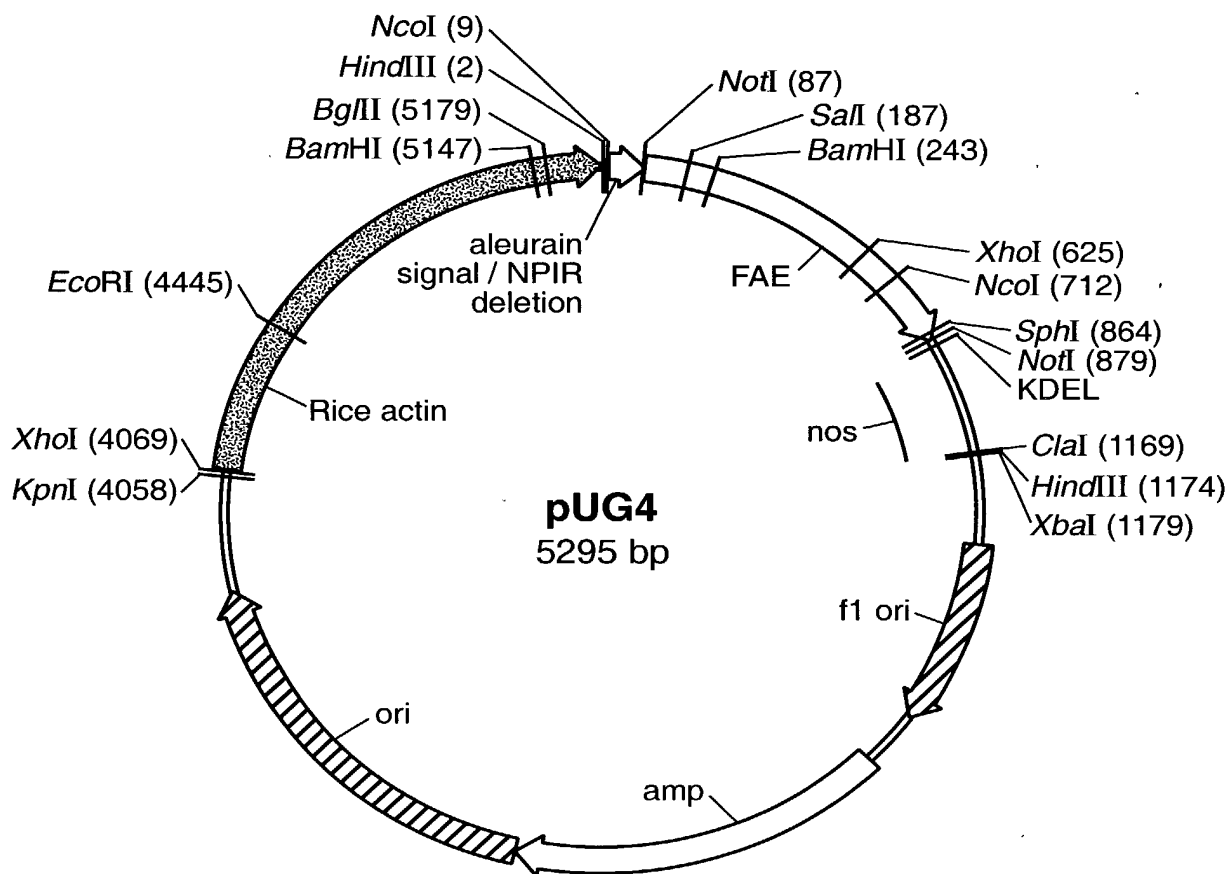
4761 TTTCCCGACT GGAAAGCGGG CAGTGAGCGC AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC
    AAAGGGCTGA CCTTTCGCCC GTCACCTCGC TTGCGTTAAT TACACTCAAT CGAGTGAGTA ATCCGTGGGG

4831 AGGCTTTACA CTTTATGCTT CCGGCTCGTA TGTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG
    TCCGAAATGT GAAATACGAA GGCCGAGCAT ACAACACACC TTAACACTCG CCTATTGTTA AAGTGTGTCC

                                BssHII                      EcoRI
                                ~~~~~
4901 AAACAGCTAT GACCATGATT ACGCCAAGCG CGCAATTAAC CCTCACTAAA GGAACAAAAA GCTGG
 TTTGTCGATA CTGGTACTAA TGCGGTTTCG GCGTTAATTG GGAGTGATTT CCTTGTGTTT CGACC
```

FIG._46H

132 / 154



**FIG._47A**

133 / 154

```

NcoI
~~~~~
HindIII
~~~~~
 M A H A R V L L L A L A V L A T A A V A V
1 AAGCTTACCA TGGCCACACG CCGCGTCTCT CTCCTGGCGC TCGCCGTGCT GCCACGGCC GCCGTGCGCG
 NotI
      ~~~~~
      . A S S R A A A S T Q G I S E D L Y S R L V E M .
71  TCGCCTCTCT CCGCGCGGCC GCCTCCACGC AGGCATCTC CGAAGACCTC TACAGCCGTT TAGTCGAAAT
      Sali
      ~~~~~
 . A T I S Q A A Y A D L C N I P S T I I K G E K
141 GGCCACTATC TCCCAAGCTG CCTACGCCGA CCTGTGCAAC ATTCCGTGCA CTATTATCAA GGGAGAGAAA
 BamHI
      ~~~~~
      I Y N S Q T D I N G W I L R D D S S K E I I T V
211  ATTTACAATT CTCAAACTGA CATTACCGA TGGATCCTCC GCGACGACAG CAGCAAAGAA ATAATCACCG
      . F R G T G S D T N L Q L D T N Y T L T P F D T .
281  TCTTCCGTGG CACTGGTAGT GATACGAATC TACAACTCGA TACTAACTAC ACCCTCACGC CTTTCGACAC
      . L P Q C N G C E V H G G Y Y I G W V S V Q D Q
351  CCTACCACAA TGCAACGGTT GTGAAGTACA CGGTGGATAT TATATTGGAT GGGTCTCCGT CCAGGACCAA
      V E S L V K Q Q V S Q Y P D Y A L T V T G H X L
421  GTCGAGTCGC TTGTCAAACA GCAGGTTAGC CAGTATCCGG ACTACGGGCT GACCGTGACC GGCCACKCCC
      . G A S L A A L T A A Q L S A T Y D N I R L Y T .
491  TCGGCGCCTC CCTGGCGGCA CTCACTGCGC CCCAGCTGTC TCGACATAC GACAACATCC GCCTGTACAC
      XhoI
      ~~~~~
 . F G E P R S G N Q A F A S Y M N D A F Q A S S
561 CTTGGCGGAA CCGCGCAGCG GCAATCAGGC CTTGCGGTCTG TACATGAACG ATGCCTTCCA AGCCTCGAGC
 P D T T Q Y F R V T H A N D G I P N L P P V E Q
631 CCAGATACGA CGCAGTATTT CCGGGTCACT CATGCCAAGC ACGGCATCCC AAACCTGCCC CCGGTGGAGC
 NcoI
      ~~~~~
```

FIG..47B

134 / 154

```

      . G Y A H G G V E Y W S V D P Y S A Q N T F V C .
701 AGGGGTACGC CCATGGCGGT GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCAGAACA CATTTGTCGTG
      . T G D E V Q C C E A Q G G Q G V N N A H T T Y
771 CACTGGGGAT GAAGTGCAGT GCTGTGAGGC CCAGGGCGGA CAGGGTGTGA ATAATGCGCA CACGACTTAT

      SphI
      ~~~~~
 F G M T S G A C T W P V A A A E P L K D E L *
841 TTTGGGATGA CGAGCGGCGC ATGCACCTGG CCGGTCGCGG CCGCGGAACC ACTGAAGGAT GAGCTGTAAA
911 GAAGCAGATC GTTCAAACAT TTGGCAATAA AGTTTCTTAA GATTGAATCC TGTGCGCGGT CTTGCGATGA
981 TTATCATATA ATTCTGTG AATTACGTTA AGCATGTAA AATTAACATG TAATGCATGA CGTTATTAT
1051 GAGATGGGTT TTTATGATTA GAGTCCCGCA ATTATACATT TAATACGCGA TAGAAAACAA AATATAGCGC

 NotI
      ~~~~~
1121 GCAAACTAGG ATAAATTATC GCGCGCGGTG TCATCTATGT TACTAGATCG ATAAGCTTCT AGAGCGGCGC
1191 GTGGAGCTCC AATTCGCCCT ATAGTGAGTC GTATTACGCG CGTCACTGG CCGTCGTTT ACAACGTCGT
1261 GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC CCTTTCGCC AGCTGGCGTA
1331 ATAGCGAAGA GGCCCGCAC GATCGCCCTT CCCAACAGTT CCGCAGCCTG AATGGCGAAT GGGACGCGCC
1401 CTGTAGCGGC GCATTAAAGC CGCGGGGTGT GGTGTTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC
1471 CTAGCGCCCG CTCCTTTCGC TTTCTTCCCT TTTCTTCTCG CCACGTTTCG CGGCTTTCCT CGTCAAGCTC
1541 TAAATCGGGG GCTCCCTTTA GGGTTCGGAT TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTA
1611 GGGTGATGGT TCACGTAGTG GGCCATCGCC CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG
1681 TTCTTTAATA GTGGACTCTT GTTCCAAACT GGAACAACAC TCAACCCCTAT CTCGGTCTAT TCTTTTGATT
1751 TATAAGGGAT TTTGCCGATT TCGGCCATT TCGTAAATAA TGAGCTGATT TAACAAAAAT TTAACGCGAA
1821 TTTTAAACAA ATATTACGC TTACAATTTA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCTATT
1891 GTTTATTTT CTAAATACAT TGAGTATTCA ACATTCCGT ATCCGCTCAT GAGACAATAA CCTGTATAA TGCCTCAATA
1961 ATATTGAAA AGGAAGAGTA TGAGTATTCA ACATTCCGT ATCCGCTCAT GAGACAATAA CCTGTATAA TGCCTCAATA
2031 TGCCTTCCGT TTTTGTCTCA CCCAGAAAAC GATCTCAACA CCGGTAAGAT CCTTGAGAGT TTTTCGCCCG AAGAACGTTT
2101 GAGTGGGTTA CATCGAACTG GATCTCAACA CCGGTAAGAT CCTTGAGAGT TTTTCGCCCG AAGAACGTTT
2171 TCCAATGATG AGCACTTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG
2241 CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACTA GAAAAGCATC
2311 TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA
2381 CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTCGC ACAACATGGG GGATCATGTA
2451 ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAAACGA CGAGCGTGAC ACCACGATGC
2521 CTGTAGCAAT GGCAACAACG TTGCGCAAAAC TATTAAGTGG CGAACTACTT ACTCTAGCTT CCGGGCAACA

```

FIG.-47C

135 / 154

```

2591 ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG
2661 TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG
2731 GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
2801 GATCGCTGAG ATAGGTGCTT CACTGATTAA GCATTGGTAA CTGTGAGACC AAGTTTACTC ATATATACTT
2871 TAGATTGATT TAAAACTTCA TTTTCTAATT AAAAGGATCT AGGTGAAGAT CCTTTTGTAT AATCTCATGA
2941 CCAAAATCCC TTAACGTGAG TTTTCTGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC
3011 TTGAGATCCT TTTTCTGTC GCGTAATCTG CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT
3081 TGTGTGCCG ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA
3151 ATACTGTCTT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT GATGACCGG GTTGGACTCA
3221 CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCTT GTCCTACCGG GTTGGACTCA
3291 AGACGATAGT TACCGGATAA GCGCAGCGG TCGGGCTGAA GCTATGAGAA AGCGCCACGC CCCAGCTTGG
3361 AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA CAGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG
3431 GAGAAAGGCG GACAGGTATC CCGTAAGCGG CAGGTCGGA ACCTCTGACT TGAGCGTCTGA TTTTGTGTAT
3501 GGAACGCCCT GGTATCTTTA TAGTCCCTGC GGGTTTCGCC ACCTCTGACT TGAGCGTCTGA TTTTGTGTAT
3571 GCTCGTCAGG GGGCGGAGC CTATGGAATA ACGCCAGCAA CGCGGCTTCT GATTCGTGG ATAACCGTAT TACCGCTTCT
3641 CTGGCCCTTT GCTCACATGT TCTTCTCTGC GTTATCCCTT GATTCGTGG ATAACCGTAT TACCGCTTCT
3711 GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACAGCCGAGC GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG
3781 AGCGCCCAAT ACGCAAAACG CCTCTCCCGG CGCGTTGGCC GATTCATTAA TGCAGCTGGC ACGACAGGTT
3851 TCCCGACTGG AAAGCGGCA GTGAGCGCAA CGCAATTAA GTGAGTTAGC TCACCTCATTA GGCACCCCCAG
3921 GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA

                                     KpnI
3991 ACAGCTATGA CCATGATTAC GCCAAGCGCG CAATTAAACC TCACTAAAGG GAACAAAAGC TGGGTACCCGG

                                     XhoI
4061 GCCCCCCTC GAGGTCAATC ATATGCTTGA GAAGAGAGTC GGGATAGTCC AAAATAAAAC AAAGGTAAGA
4131 TTACCTGGTC AAAAGTGAA ACATCAGTTA AAAGGTGTA TAAGTAAAT ATCGGTAATA AAAGGTGGCC
4201 CAAAGTGAAA TTTACTCTTT TCTACTATTA TAAAAATTGA GGATGTTTGG TCGGTACTTT GATACGTCTAT
4271 TTTTGTATGA ATTGGTTTTT AAGTTTATTC GCGATTGGA AATGCATATC TGTATTTGAG TCGGTTTTTA
4341 AGTTCGTTGC TTTTGTAAAT ACAGAGGGAT TTGTATAAGA AATATCTTTA AAAAACCCCAT ATGCTAATTT

                                     EcoRI
4411 GACATAAATT TTGAGAAAAA TATATATTCA GGCGAATTCC ACAATGAACA ATAATAAGAT TAAAAATAGCT
4481 TGCCCCCGTT GCAGCGATGG GTATTTTCTC TAGTAAAAA AAAGATAAAC TTAGACTCAA AACATTTACA
4551 AAAACAACCC CTAAAGTCTT AAAGCCCAA GTGCTATGCA CGATCCATAG CAAGCCCAAG CCAACCCCAAC
4621 CCAACCCCAAC CCACCCCAAGT GCAGCCCACT GGCAAAATAGT CTCCACCCCC CCGTACTATCA CCGTGAGTTG

```

FIG.- 47D

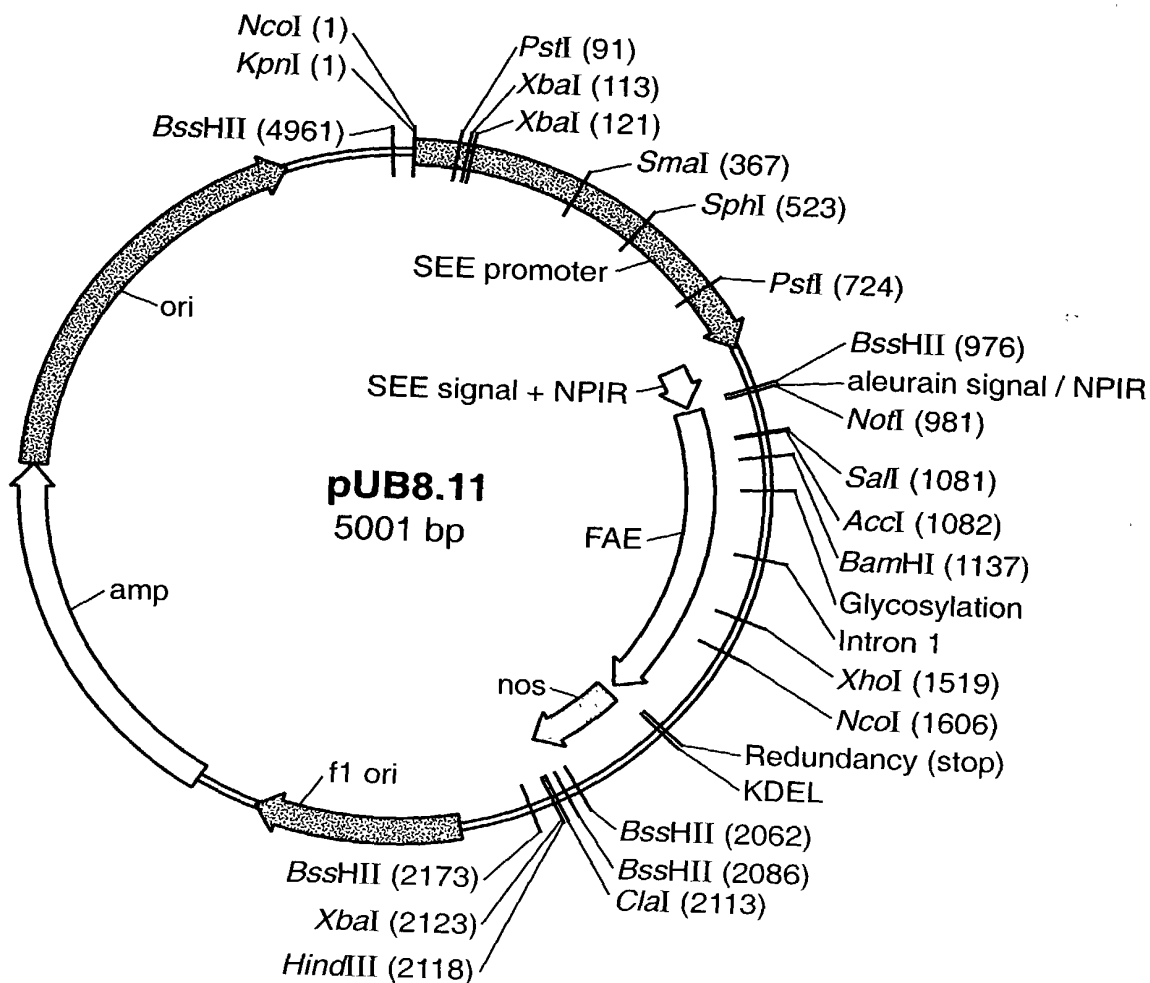
136 / 154

```
4691 TCCGCACCCAC CGCACGTCTC GCAGCCAAA AAAAAGAG AAGAAAAAGA AAAACAGCAG
4761 GTGGGTCCGG GTCGTGGGG CCGGAAAAGC GAGGAGATC GCGAGCAGCG ACGAGGCCCG GCCCTCCCCTC
4831 CGCTTCCAAA GAAACGCCCC CCATCGCCAC TATATACATA CCCCCCCTC TCCTCCCATC CCCCCAACCC
4901 TACCACCACC ACCACCACCA CTTCTTCTCT CTTCTTCTCT CTTCTTCTCT CTTCTTCTCT CTTCTTCTCT
4971 GCCGCCGCCG GTAACCAACC CGCCCTCTCT GTGGGCGAGA GCGGCTTCTG CGCCCAAGATC GGTGCGCGGG AGGGCGCGGA
5041 CTTTGGCCCTT GGTAGTTTGG GTGGGCGAGA GCGGCTTCTG CGCCCAAGATC GGTGCGCGGG AGGGCGCGGA
                                         BamHI
                                         ~~~~~~
5111 TCTCGCGGCT GCGGTCTCCG GCGGTGAGTC GGCCCGGATC CTCGCGGGGA ATGGGGCTCT CGGATGTAGA
 BglII
                                         ~~~~
5181 TCTTCTTTCT TTTCTTTTCT TGTGGTAGAA TTTGAATCCC TCAGCATTTG TCATCGGTAG TTTTTCCTTT
5251 CATGATTTGT GACAAATGCA GCCTCGTGCG GAGCTTTTTT GTAGC
```

FIG..47E



137 / 154



**FIG. 48A**

138 / 154

```

NcoI
~~~~~
KpnI
~
1 CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
 GTACCCGGTC CATATTAATA CCCATAGAG TTCGTTTATT AGCTTTATAG TGGTAACCGA TGTATAGAC

 PstI
                                     ~~~~~
                                     XbaI
                                     ~~~~~
71 AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTGTG ATCTAGAACT CTAGATAGCA CAGCCACAGC
 TCGAGGCTCA AGACTGACGT CAGACCTACT CAGCACAACA TAGATCTTGA GATCTATCGT GTCGGTGTCTG

 XbaI
                                     ~~~~~
141 ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTG GAGACGGAGC TCTTTCCTAC CTCTTGACGT
   TGGATGTCTT CACGCTGTGA ACACCTGACA TCATCACAAC CTCTGCCCTCG AGAAAGGATG GAGGACTGCA

211 TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC CCAACAAAAT ATCGTCCCCC
   ACGCGGGCAA CAGGTAAGGT TGCCGTAGTG AGAGTTGGTT AGTGCGCGAG GGTGTGTTTA TAGCAGGGGG

281 ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTGCGGCC GTTTTGTCT GAATCTCGCT TCCACTGGCC
   TACAGAACCG CCTCTCTCTC ATGTATGTAC GACAGCGCG CAAAACAGA CTTAGAGCGA AGGTGACCCG

                                     SmaI
                                     ~~~~~
351 AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCTG TCACCCCTGG CGTCATGGGA
 TTAGTCGAGT CGAGGGCCCT CGAGTGAGTA AGTTCTAGG TAGCAGCAGC AGTGGGGACC GCAGTACCCT

421 TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAACCCCAAT
 ACCTTTTCTT GGAGGCAACG AGCCTACTCA GTCGGTATAG GGGCTTGTCT CATGACGTTT TATTGGGTTA

 SphI
                                     ~~~~~
491 TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTGTTG CTTAATTGAC TTTATTTTGT
   AGTCTAAGGG GGTATATCTCT TTCAATATCGT ACGAAAGCCC AAAACAAACC GAATTAACCTG AAATAAAAC

561 TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT CGAGACGGAT AATAGGCTGG
   AACCTCAACT TACGACTAAA CAACACATTT TACGGGTTGG TAGACTTATA GCTCTGCCTA TTATCCGACC
  
```

FIG..48B

139 / 154

```

631  CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGGCATTTACA GCTGGAGGCT
    GATTAATTAA ATATCGTTCT AAGACATCAC GTGTAGCGTT TATAGAAAGA CCCGTAATGT CGACCTCCGA

          PstI
          ~~~~~
701 TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG CGATGAGATG GGTATAAAAC
 AGTAGTCGGA CTTTGTGAGA CGTCTCGGAC TTTCGTCACC ACTTCGCACC GCTACTCTAC CCATATTTTG

771 CCCGGGCACC GGGACGCGAG CTCCC GCCCTA CCAGTACCAT CTCGCCCTCGC TCCCCCTGCC GGACGACCCA
 GGGGCCGTGG CCTGCGCTC GAGGCGGAT GGTCAATGTA GAGCGGAGCG AGGGGACGG CCTGCTGGGT

841 GTAAATACT GTTGCCCACT CGCCGGCGAG ATGGCCACG GCCGCATCCT CTTCTTTGGCG CTCGCCGTCT
 CATTTATGA CAACGGGTGA GCGGCCGCTC TACCGGGTGC CGGCGTAGGA GAAGAACCGC GAGCGGCAGA

 BssHII
          ~~~~~
          NotI
          ~~~
911 TGGCCACCGC CGCGGTGGCC GCCGCATCMT TGGCGGACTC CAACCCGATC CGGCCCGTCA CCGAGCGCGC
 ACCGGTGGCG GCGCCACCGG CGCGGTAGNA ACCGCTGAG GTTGGGCTAG GCCGGGCAGT GGCTCGCGCG

 NotI
          ~~~~~
981  GGCCGCCTCC ACGCAGGGCA TCTCCGAAGA CCTCTACAGC CGTTTAGTCG AAATGGCCAC TATCTCCCAA
    CCGGCGGAGG TGCCTCCCGT AGAGGCTTCT GGAGATGTCG GCAAAATCAGC TTACCCGGTG ATAGAGGTTT

          Sali
          ~~~~~
 AccI
          ~~~~~
1051 GCTGCCCTACG CCGACCTGTG CAACATTCCG TCGACTATTA TCAAGGGAGA GAAAATTAC AATCTCAA
    CGACGGATGC GGCTGGACAC GTTGTAAGGC AGCTGATAAT AGTTCCCTCT CTTTAAATG TTAAGAGTTT
```

FIG._48C

140 / 154

```

      BamHI
      ~~~~~
1121 CTGACATTAA CGGATGGATC CTCGCGGACG ACAGCAGCAA AGAAATAATC ACCGTCCTCC GTGGCACTGG
 GACTGTAATT GCCTACCTAG GAGGCGCTGC TGTCGTCGTT TCTTTATTAG TGGCAGAAGG CACCGTGACC

1191 TAGTGATACG AATCTACAAC TCGATACTAA CTACACCCCTC ACGCCTTTTCG ACACCCCTACC ACAATGCAAC
 ATCACTATGC TTAGATGTTG AGCTATGATT GATGTGGGAG TGCAGAAAGC TGTGGGATGG TGTACGTTG

1261 GGTGTGAAG TACACGGTGG ATATTATATT GGATGGGTCT CCGTCCAGGA CCAAGTCGAG TCGCTTGTCA
 CCAACACTTC ATGTGCCACC TATAATATAA CCTACCCAGA GGCAGGTCCT GGTTCAGCTC AGCGAACAGT

1331 AACAGCAGGT TAGCCAGTAT CCGGACTACG CGCTGACCGT GACCGGCCAC KCCCTCGGCG CCTCCCTGGC
 TTGTCGTCCA ATCGGTCATA GGCCTGATGC GCGACTGGCA CTGGCCGGTG MGGGAGCCGC GGAGGGACCG

1401 GGCACCTCACT GCCGCCCAGC TGTCTGCGAC ATACGACAAC ATCCGCCCTGT ACACCTTCGG CGAACC CGCG
 CCGTGAGTGA CGGCGGGTCG ACAGACGCTG TATGCTGTTG TAGGCGGACA TGTGGAAGCC GCTTGCGCGG

 XhoI
      ~~~~~
1471 AGCGGCAATC AGGCCTTTCG GTCGTACATG AACGATGCCT TCCAAGCCTC GAGCCCAGAT ACGACGCAGT
      TCGCCGTTAG TCCGGAAGCG CAGCATGTAC TTGCTACGGA AGGTTCCGAG CTCGGGTCTA TGCTGCGTCA

      NcoI
      ~~~~~
1541 ATTTCCGGGT CACTCATGCC AACGACGGCA TCCCAAACCT GCCCCCGGTG GAGCAGGGGT ACGCCCCATGG
 TAAAGGCCCA GTGAGTACGG TTGCTGCCGT AGGGTTTGA CGGGGCCAC CTCGTCCCCA TGCGGGTACC

1611 CGGTGTAGAG TACTGGAGCG TTGATCCTTA CAGCGCCCAG AACACATTG TCTGCACCTGG GGATGAAGTG
 GCCACATCTC ATGACCTCGC AACTAGGAAT GTCGCGGGTC TTGTGTAAC AGACGTGACC CCTACTTCAC

1681 CAGTGTCTGT AGGCCCCAGG CGGACAGGGT GTGAATAATG CGCACACGAC TTATTTTGGG ATGACGAGCG
 GTCACGACAC TCCGGGTCCC GCCTGTCCCA CACTTATTAC GCGTGTGCTG AATAAAACCC TACTGCTCGC

1751 GAGCCTGTAC ATGGTGATCA GTCATTTTCAG CCTCCCCGAG TGTACCAGGA AAGATGGATG TCCTGGAGAG
 CTCGGACATG TACCACTAGT CAGTAAAGTC GGAGGGGCTC ACATGGTCCT TTCTACCTAC AGGACCTCTC

```

FIG. 48D

141 / 154

```

1821 GGGCCGCGT AACCACTGAA GGATGAGCTG TAAAGAAGCA GATCGTTCAA ACATTGGCA ATAAAGTTTC
 CCCCGGCGCA TTGGTGACTT CCTACTCGAC ATTCTCTCGT CTAGCAAGTT TGTAACCCGT TATTTCAAAG

1891 TTAAGATTGA ATCCTGTTGC CGGTCTTGCG ATGATTATCA TATAATTCTT GTTGAATTAC GTTAAGCATG
 AATCTTAAT TAGGACAACG GCCAGAACGC TACTAATAGT ATATTAAAGA CAACCTTAATG CAATTCGTAC

1961 TAAATAATTA CATGTAATGC ATGACGTTAT TTATGAGATG GGTTTTTATG ATTAGAGTCC CGCAATTATA
 ATTATTAATT GTACATTACG TACTGCAATA AATACTCTAC CCAAAAATAC TAATCTCAGG GCGTTAATAT

 BssHII
      ~~~~~
2031 CATTTAATAC GCGATAGAAA ACAAAATATA GCGCGCAAC TAGGATAAAT TATCGCGCGC GGTGTCATCT
      GTAAATTATG CGCTATCTTT TGTTTTATAT CGCGCGTTTG ATCCTATTTA ATAGCGCGCG CCACAGTAGA

      XbaI
      ~~~~~
 ClaI HindIII
      ~~~~~
2101 ATGTTACTAG ATCGATAAGC TTCTAGAGCG GCCGGTGGAG CTCCAATTGC CCCATATAGT AGTCGTATTA
      TACAATGATC TAGCTATTGC AAGATCTCGC CGGCCACCTC GAGGTTAAGC GGGATATCAC TCAGCATAAT

      BssHII
      ~~~~~
2171 CGCGCGCTCA CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACCCCTG GCGTTACCCA ACTTAATCGC
 GCGCGCGAGT GACCGGCAGC AAAATGTTGC AGCACTGACC CTTTTGGGAC CGCAATGGGT TGAATTAGCG

2241 CTTGCAGCAC ATCCCCCTTT CGCCAGCTGG CGTAATAGCG AAGAGGCCCG CACCGATCGC CCTTCCCAAC
 GAACGTCGTG TAGGGGAAA GCGGTCGACC GCATTATCGC TTCTCCGGGC GTGGCTAGCG GGAAGGGTTG

2311 AGTTGCGCAG CCTGAATGGC GAATGGGACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG GTGTGGTGGT
 TCAACGCGTC GGACTTACCG CTTACCCCTG CCGGGACATC GCCGCGTAAT TCGCGCCGCC CACACCACCA

2381 TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCTTT TCGCTTTCTT CCCTTCCTTT
 ATGCGCGTCG CACTGGCGAT GTGAACGGTC GCGGGATCGC GGGCGAGGAA AGCGAAAGAA GGAAGGAAA

```

FIG..48E

2451 CTCGCCACGT TCGCCGGGCTT TCCCCGTCAG GCTCTAAATC GGGGGCTCCC TTTAGGGTTC CGATTAGTGC  
 GAGCGGTGCA AGCGGCCGAA AGGGCAGTT CGAGATTAG CCCCCGAGG AAATCCCAAG GCTAAATCAC

2521 CTTTACGGCA CCTCGACCCC AAAAAACTTG ATTAGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA  
 GAAATGCCGT GGAGCTGGG TTTTGTGAAC TAATCCCACT ACCAAGTGCA TCACCCCGGTA GCGGGACTAT

2591 GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCCTA AACTGGAACA  
 CTGCCAAAAA GCGGGAAACT GCAACCTCAG GTGCAAGAAA TTATCACCTG AGAACCAAGT TTGACCTTGT

2661 ACACCTCAACC CTATCTCGGT CTATCTCTTT GATTATAAG GGATTTTCCC GATTTCCGCC TATTGGTTAA  
 TGTGAGTTGG GATAGAGCCA GATAAGAAAA CTAAATATTCT CTTAAACCGG CTAAAGCCGG ATAACCAATT

2731 AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTTAA CAAAATATTA ACGCTTACAA TTTAGGTGGC  
 TTTTACTCGA CTAAATTGTT TTTAAATTGC GCTTAAATTT GTTTATAAT TGCGAATGTT AAATCCACCG

2801 ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT ACATTCAAAT ATGTATCCGC  
 TGAAAAAGCC CTTTACACGC GCCTTGGGGA TAAACAAATA AAAAGATTTA TGTAAGTTTA TACATAGGCG

2871 TCATGAGACA ATAAACCCTGA TAAATGCTTC AATAATATTG AAAAAAGGAA AGTATGAGTA TTCAACATTT  
 AGTACTCTGT TATTGGGACT ATTTACGAAG TTATTATAAC TTTTCTCTTC TCATACTCAT AAGTTGTAAA

2941 CCGTGTCGCC CTTATTCCCT TTTTTCGGC ATTTTCCTT CCTGTCTTTG CTCACCCAGA AACGCTGGTG  
 GGCACAGCGG GAAATAAGGA AAAAACGCCG TAAACCGGAA GGACAAAAAC GAGTGGGTCT TTGCGACCAC

3011 AAAGTAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA  
 TTTCAATTTT TACGACTTCT AGTCAACCCA CGTGCTCACC CAATGTAGCT TGACCTAGAG TTGTGCGCCAT

3081 AGATCCTTGA GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTTAAAGTTC TGCTATGTGG  
 TCTAGGAACT CTCAAAAGCG GGGCTTCTTG CAAAAGTTA CTACTCGTGA AAATTTCAAG ACGATACACC

3151 CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC GGTCCGCCGA TACACTATTC TCAGAATGAC  
 GCGCCATAAT AGGCGATAAC TGCGGCCCGT TCTCGTTGAG CCAGCGGCGT ATGTGATAAG AGTCTTACTG

3221 TTGGTTGAGT ACTCACCAGT CACAGAAAAA CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG  
 AACCAACTCA TGAGTGGTCA GTGTCTTTTC GTAGAATGCC TACCGTACTG TCATTCTCTT AATACGTCAC

FIG. 48F

143 / 154

3291 CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC CGAAGGAGCT  
GACGGTATTG GTACTCACTA TTGTGACGCC GGTGAATGA AGACTGTTGC TAGCCTCCTG GCTTCCTCGA

3361 AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACTCGC CTTGATCGTT GGGAAACCGGA GCTGAATGAA  
TTGGCGAAAA AACGTGTTGT ACCCCCTAGT ACATTGAGCG GAACCTAGCAA CCCTTGGCCT CGACTTACTT

3431 GCCATACCAA ACGACGAGCG TGACACCACG ATGCCCTGTAG CAATGGCAAC AACGTTGCGC AAACCTATTAA  
CGGTATGGTT TGCTGCTCGC ACTGTGGTGC TACGGACATC GTTACCCTGTG TTGCAACGCG TTTGATAAAT

3501 CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG  
GACCGCTTGA TGAATGAGAT CGAAGGGCCG TTGTTAATTA TCTGACCTAC CTCCGCCCTAT TTCAACGCTCC

3571 ACCACTTCTG CGCTCGGCC TTCCGGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG  
TGGTGAAGAC GCGAGCCGGG AAGGCCGACC GACCAAAATA CGACTATTTA GACCTCGGCC ACTCGCACCC

3641 TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG  
AGAGCGCCAT AGTAACGTCTG TGACCCCGGT CTACCATTCTG GGAGGGCATA GCATCAATAG ATGTGCTGCC

3711 GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG  
CCTCAGTCCG TTGATACCTA CTTGCTTTAT CTGTCTAGCG ACTCTATCCA CGGAGTGACT AATTCGTAAAC

3781 GTAACTGTCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAC TTCATTTTTTA ATTTAAAGG  
CATTGACAGT CTGGTTCAAA TGAGTATATA TGAAATCTAA CTAATTTTG AGTAAAAAT TAAATTTTCC

3851 ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAGAA TCCCTTAACG TGAGTTTTCG TTCCACTGAG  
TAGATCCACT TCTAGGAAA ACTATTAGAG TACTGGTTTT AGGGAATTGC ACTCAAAAGC AAGGTGACTC

3921 CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT  
GCAGTCTGGG GCATCTTTTC TAGTTTCCCTA GAAGAACTCT AGGAAAAAAA GACGCGCATT AGACGACGAA

3991 GCAACAAAA AAACCACCGC TACCAGCGGT GGTGTTGTTG CCGGATCAAG AGCTACCAAC TCTTTTTCG  
CGTTTGTGTTT TTTGGTGGCG ATGGTCGCCA CCAACAAAA GGCCTAGTTC TCGATGGTTG AGAAAAAGG

4061 AAGGTAACCTG GCTTCAGCAG AGCGCAGATA CCAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC  
TTCCATTGAC CGAAGTCGTC TCGCGTCTAT GGTATTGAC AGGAAGATCA CATCGGCATC AATCCGGTGG

FIG._48G

144 / 154

4131 ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG TTACCAGTGG CTGCTGCCAG  
TGAAGTTCTT GAGACATCGT GCGGGATGTA TGGAGCGAGA CGATTAGGAC AATGGTCACC GACGACGGTC

4201 TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTCGGGC  
ACCGCTATTG AGCACAGAAT GGCCCAACCT GAGTTCTGCT ATCAATGGCC TATTCCGCGT CGCCAGCCCG

4271 TGAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC  
ACTTGCCCCC CAAGCACGTG TGTCGGGTG AACCTCGCTT GCTGGATGTG GCTTGACTCT ATGGATGTCG

4341 GTGAGCTATG AGAAAGCGCC ACGCTTCCG AAGGGAGAAA GGCGGACAGG TATCCGGTAA GCGGCAGGGT  
CACTCGATAC TCTTTCGCGG TCGGAAGGGC TTCCCTCTTT CCGCCTGTCC ATAGGCCATT CGCCGTCCCA

4411 CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT  
GCCTTGTCTT CTCGCGTGCT CCTCGAAGG TCCCCCTTTG CGGACCATAG AAATATCAGG ACAGCCCAAA

4481 CGCCACCTCT GACTTGAGCG TCGATTTTGT TGAATGCTCGT CAGGGGGGCG GAGCCTATGG AAAAACGCCA  
GCGGTGGAGA CTGAACCTCG AGCTAAAAAC ACTACGAGCA GTCCCCCCCG CTCGGATACC TTTTTCGGT

4551 GCAACGCGGC CTTTTTACGG TTCTTGGCCT TTTTGCTGCC TTTTGCTCAC ATGTTCTTTC CTGCGTTATC  
CGTTGCGCCG GAAAAATGCC AAGGACCGGA AAACGACCGG AAAACGAGTG TACAAGAAAAG GACGCAATAG

4621 CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACCG CTCGCCCGCAG CCGAACGACC  
GGGACTAAGA CACCTATTGG CATAATGGCG GAAACTCACT CGACTATGGC GAGCGGCGTC GGCTTGCTGG

4691 GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA ACCGCCCTCTC CCCGCGCGTT  
CTCGCGTCGC TCAGTCACTC GCTCCTTCGC CTTCTCGCGG GTTATGCGTT TGGCGGAGAG GGGCGCGCAA

4761 GGCCGATTCA TTAATGCAGC TGGCACGACA GGTTCCCGA CTGGAAGCG GGCAGTGAGC GCAACGCAAT  
CCGGCTAAGT AATTACGTCG ACCGTGCTGT CCAAAGGGCT GACCTTTCGC CCGTCACTCG CGTTGCGTTA

4831 TAAATGTGAG TAGCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC TTCCGGCTCG TATGTTGTGT  
ATTACACTCA ATCGAGTGAG TAATCCGTGG GGTCCGAAAT GTGAAATACG AAGGCCGAGC ATACAACACA

FIG. 48H



4901 GGAATTGTGA GCGGATAACA ATTTACACACA GGAACACAGCT ATGACCATGA TTACGCCAAG CGCGCAATTA  
CCTTAACACT CGCCTATTGT TAAAGTGTGT CCTTTGTCGA TACTGGTACT AATGCGGTTC GCGCGTTAAT

BssHI  
~~~~~

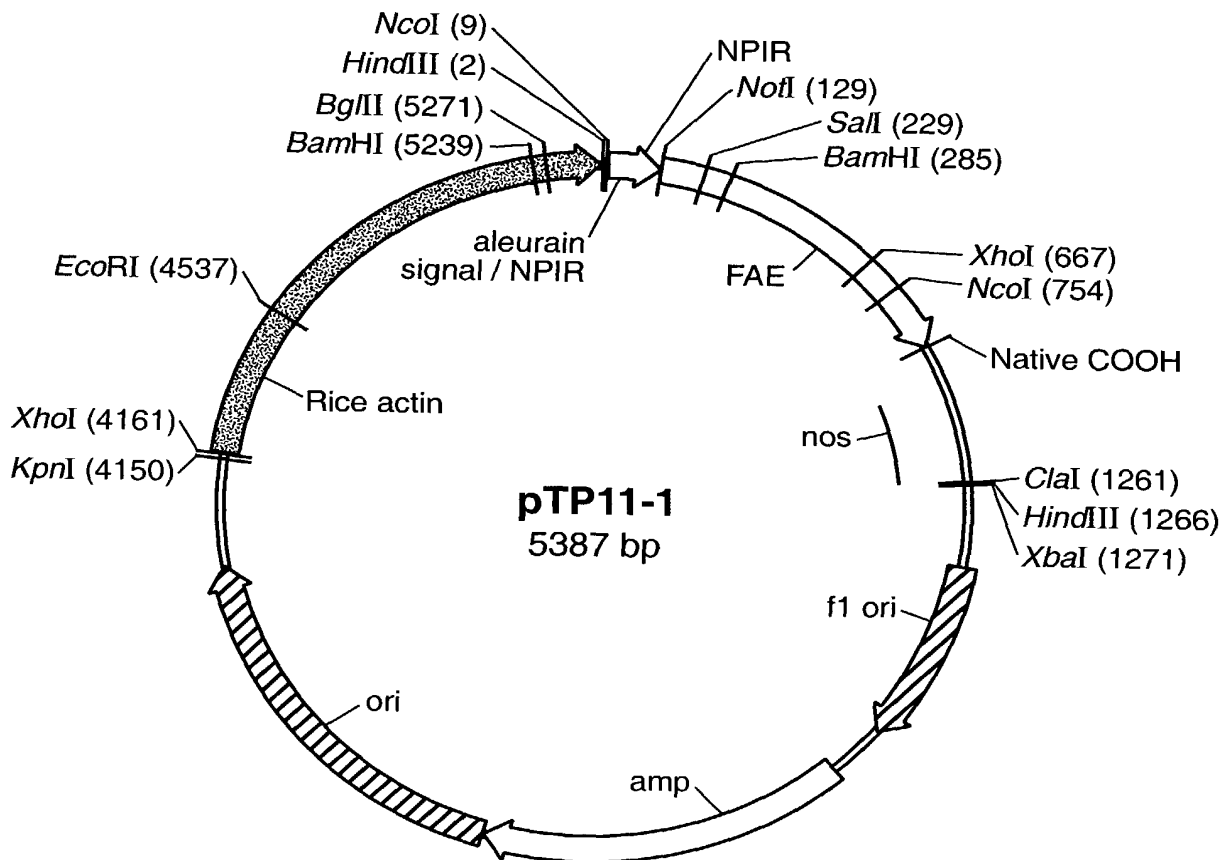
4971 ACCCTCACTA AAGGGAACAA AAGCTGGGTA C  
TGGGAGTGAT TTCCCTTGTT TTCGACCCAT G

NcoI  
~~~~~

KpnI  
~~~~~

FIG.\_48I

146 / 154



**FIG..49A**

147 / 154

```

 NcoI
      ~~~~~
HindIII
~~~~~
 M A H A R V L L L A L A V L A T A A V A V
1 AAGCTTACCA TGGCCACGC CCGGTCTCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTGCGCC
 NotI
      ~~~~~
      . A S S S S F A D S N P I R P V T D R A A A S T .
71 TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCGATCCG GCCGTCACC GACCGCGCGG CCGCCTCCAC
      . Q G I S E D L Y S R L V E M A T I S Q A A Y A
141 GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC
      SalI
      ~~~~~
      ~~~~~
      AccI
      ~~~~~
 D L C N I P S T I I K G E K I Y N S Q T D I N G
211 GACCTGTGCA ACATTCCGTC GACTATTATC AAGGAGAGA AAATTACAA TTCTCAAACT GACATTACG
 BamHI
      ~~~~~
      . W I L R D D S S K E I I T V F R G T G S D T N .
281 GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACGTGTA GTGATACGAA
      . L Q L D T N Y T L T P F D T L P Q C N G C E V
351 TCTACAACTC GATACTAACT ACACCTTCAC GCCTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
      H G G Y Y I G W V S V Q D Q V E S L V K Q Q V S
421 CACGGTGGAT ATTATATTGG ATGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAA CAGCAGGTTA
      . Q Y P D Y A L T V T G H X L G A S L A A L T A .
491 GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
      . A Q L S A T Y D N I R L Y T F G E P R S G N Q
561 CGCCGAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
      XhoI
      ~~~~~
 A F A S Y M N D A F Q A S S P D T T Q Y F R V T
631 GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA

```

FIG. 49B

148 / 154

```

 H A N D G I P N L P P V E Q G Y A H G G V E Y .
701 CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGA GCAGGGGTAC GCCATGGCG GTGTAGAGTA
 W S V D P Y S A Q N T F V C T G D E V Q C C E
771 CTGGAGCGTT GATCCTTACA GCGCCAGAA CACATTGTC TGCACCTGGG ATGAAGTGCA GTGCTGTGAG
 A Q G G Q G V N N A H T T Y F G M T S G A C T W
841 GCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTGGGAT GACGAGCGGA GCCTGTACAT
 *
911 GGTGATCAGT CATTTCAGCC TCCCCGAGTG TACCAGGAA GATGGATGTC CTGGAGAGGG GGCCGCGTAA
981 CCACTGAAGG ATGAGCTGTA AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT AAAGTTTCTT AAGATTGAAT
1051 CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTCTGT TGAATTACGT TAAGCATGTA ATAATTAAACA
1121 TGTAATGCAT GACGTTATTT ATGAGATGGG TTTTATGAT TAGAGTCCCG CAATTATACA TTTAATACGC
 ClaI
1191 GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG TGTCACTCTAT GTTACTAGAT
 HindIII
 ClaI XbaI
      ~~~~~
1261 CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC CTATAGTGAG TCGTATTACG CGCGTCACT
1331 GGCGTTCGTT TTACAACGTC GTGACTGGGA AAACCTTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT
1401 CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC
1471 TGAATGGCGA ATGGGACGCG CCCTGTCGCG CGCTCCTTTC TAGGGTTCCG ATTTAGTGCT CGCCACGTTT
1541 GACCGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTTCCG CCTGATAGA CGGTTTTCG
1611 TCGACCCCAA AAACTTGAT TAGGGTGATG GTTCCTTAA TAGTGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCCCT
1751 CCCTTTGACG TTGGAGTCCA ATCTTTTGA TTTATAAGGG ATTTTGCCGA TTTTCGGCTA TTGGTTAAAA AATGAGCTGA
1821 ATCTCGGTCT ATCTTTTGA TTTATAAGGG ATTTTGCCGA TTTTCGGCTA TTGGTTAAAA AATGAGCTGA
1891 TTTAACAAAA ATTTAACGCG AATTTAACA AAATATTAAC GCTTACAATT TAGGTGGCAC TTTTCGGGGA
1961 AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTTCAAATAT GTATCCGCTC ATGAGACAAT
2031 AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCG GTGTCGCCCT
2101 TATTCCCTTT TTTGCGGCAT TTTGCCCTTC TGTTTTGTCT CACCCAGAAA CGCTGCTGAA AGTAAAAGAT
2171 GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA
2241 GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC
2311 CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCGATA CACTATTCTC AGAATGACTT GGTGAGTAC
2381 TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAAAT ATGCAGTGCT GCCATAACCA

```

FIG. 49C

```
2451 TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT
2521 GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAAC
2591 GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCGAACTAC
2661 TTAATCTAGC TTCCCGGCCA CAATTAATAG ACTGGATGGA GCGGATATAA GTTGACGGAC CACTTCTGCG
2731 CTCGGCCCTT CCGCTGGCT TGGTAAATTC TGGTAAATTC TCCCGTATCG TAGTTATCTA CACGACGGG AGTCAGGCAA
2801 ATTGCAGCAC TGGGGCCAGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACGTGAT TTAAGTGGT AACTGTCAGA
2871 CTATGGATGA ACATGATATG TTTAGATTGA TTTTAAACTT CATTTTAACT TTAAGTGGT CTAGGTGAAG
2941 CCAAGTTTAC TCATATATAC TTTAGATTGA TTTTAAACTT CATTTTAACT TTAAGTGGT CTAGGTGAAG
3011 ATCCTTTTTG ATAACTCTAT CAAGGATCTT TCTTGAGATC CTTTGTTCCT GCGGTAATC TGCTGCTTGC AAACAAAAA
3081 TAGAAAAGAT CAAAGGATCTT CAAGGATCTT TCTTGAGATC CTTTGTTCCT GCGGTAATC TGCTGCTTGC AAACAAAAA
3151 ACCACCGCTA CCAGCGGTGG TTTGTTTGGC GGATCAAGAG GTTACCAACTC TTTTTCGAA GGTAACTGGC
3221 TTCAGCAGAG CGCAGATACC GCCTACATAC AAATACTGTC CTTCTAGTGT AGCGTAGTT AGCCACCCAC TTCAAGAACT
3291 CTGTAGCACC GCCTACATAC CTTGCTCTGC TAATCTGTGT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC
3361 GTGTCTTACC GGTGTTGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTGCGGCTG AACGGGGGT
3431 TCGTGACACAC AGCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG
3501 AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CCGACAGGTA TCCGGTAAAG GGCAGGTCG GAACAGGAGA
3571 GCGCACGAGG GAGCTTCCAG GAGCTTCCAG GGGGAGGCGA GCTATGGA GCCTATGGA AACGCCAGC AACCGGCCCT
3641 CTTGAGCGTC GATTTTGTG ATGCTCGTCA GGGGCGCGA TTGCTCACAT GTTCTTTCCT CGTTATCCC CTGATTTCTGT
3711 TTTTACGGTT CTTGCGCTTT ATTACCGCTT TTGAGTGAGC TGATACCGCT CGCGCAGCC GACGACGAG
3781 GGATAACCGT ATTACCGCTT TTGAGTGAGC TGATACCGCT CGCGCAGCC GACGACGAG CCGATTCATT
3851 TCAGTGAGCG AGGAAGCGGA AGAGCGCCA ATACGCAAC GAAAGCGGG CAGTGAGCGC CAGTGAAGTA
3921 AATGACGCTG GCACGACAGG TTTCCCGACT GAAAGCGGG CAGTGAAGCG CAGTGAAGCG CAGTGAAGCG
3991 GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT GAAAGCGGG CAGTGAAGCG CAGTGAAGCG
4061 GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT ACGCCAAAGC CGCAATTAAC CCTCACTAAA

KpnI
~~~~~
4131 GGGAAACAAA GCTGGGTACC GGGCCCCCCC TCGAGGTCAT TCATATGCTT GAGAAAGAGAG TCGGGATAGT
4201 CCAAAATAAA ACAAGGTAA GATTACCTGG TCAGAAAGTGA AAACATCAGT TAAAAGGTGG TATAAGTAAA
4271 ATATCGGTAA TAAAAGGTGG CCCAAAGTGA AATTACTCT TTTCTACTAT TATAAAAATT GAGGATGTTT
4341 TGTCGGTACT TTGATACGTC ATTTTGTAT GAATTGGTTT TTAAGTTTAT TCGCGATTTG GAAATGCATA
4411 TCTGTATTG AGTCGGTTT TAAAGTTCGT GCTTTGTAA ATACAGAGG ATTGTATATA GAAATATCTT

EcoRI
~~~~~
4481 TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA AATATATATT CAGGCGAATT CCACAATGAA
4551 CAATAATAAG ATTAAAAATG CTTGCCCCCG TTGCAGCGAT GGGTATTTT TCTAGTAAAA TAAAAAGATAA
4621 ACTTAGACTC AAAACATTA CAAAAACAAC CCTAAAGTC CTAAAGCCCA AAGTGTATG CACGATCCAT
```

FIG. 49D

150 / 154

```
4691 AGCAAGCCCA GCCCAACCCA ACCCAACCCA ACCCAACCCA GTGCAGCCAA CTGGCAATA GTCTCCACCC
4761 CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC TCGCAGCCAA AAAAAAAAAA AGAAGAAAA
4831 AAAAGAAAA GAAAACAGC AGGTGGGTCC GGTGCGTGGG GGCCGGAATA GCGAGGAGGA TCGCGAGCAG
4901 CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC CCCATCGCC ACTATATACA TACCCCCCCC
4971 TCTCCTCCCA TCCCCCAAC CCTACCACCA CCACCACCAC CACCTCCTCC CCCCTCGCTG CCGGACGACG
5041 AGCTCCTCCC CCTCCCCCT CCGCCGCCGC CGTAACCAAC CCCGCCCCTC TCCTCTTTCT TTCTCCGTTT
5111 TTTTTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT GGTGGGCCGA GAGCGGCTTC GTCGCCCCAGA

                                     BamHI
                                     ~~~~~~

5181 TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC CGGGCGGTGAG TCGGCCCCGA TCCTCGCGGG

 BglII
                                     ~~~~~~

5251 GAATGGGGCT CTCGGATGTA GATCTTCTTT CTCTCTCTTT TTTGTGGTAG AATTGAATC CCTCAGCATT
5321 GTTCATCGGT AGTTTTTCTT TTCATGATTT GTGACAAATG CAGCCTCGTG CGGAGCTTTT TTGTAGC
```

FIG.- 49E

**Actin Promoter - FAEs**

Diagram illustrating the Actin Promoter - FAEs construct. The construct is 1259 bp long. Key features include the FULL KpnI-EcoRI deletion, and restriction sites for KpnI (6), XhoI (17), EcoRI (393), HindIII (1248), pCOR deletion, BglII (1127), and BamHI (1095).

**ACTIN-PROMOTER-FAEs**  
1259 bp

|     | KpnI              | XhoI              |                    |                   |                   |                   |                   |
|-----|-------------------|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|
|     | ~~~~~             | ~~~~~             |                    |                   |                   |                   |                   |
| 1   | <u>GGTACCGGGC</u> | <u>CCCCCCTCGA</u> | <u>GGTCATTTCAT</u> | <u>ATGCTTGAGA</u> | <u>AGAGAGTCGG</u> | <u>GATAGTCCAA</u> | <u>AATAAAACAA</u> |
|     | <u>CCATGGCCCG</u> | <u>GGGGGAGCT</u>  | <u>CCAGTAAGTA</u>  | <u>TACGAACTCT</u> | <u>TCTCTCAGCC</u> | <u>CTATCAGGTT</u> | <u>TTATTTTGTT</u> |
| 71  | <u>AGGTAAGATT</u> | <u>ACCTGGTCAA</u> | <u>AAGTGAAAC</u>   | <u>ATCAGTTAAA</u> | <u>AGGTGGTATA</u> | <u>AGTAAATAT</u>  | <u>CGGTAATAAA</u> |
|     | <u>TCCATTCTAA</u> | <u>TGGACCAGTT</u> | <u>TTCACTTTTG</u>  | <u>TAGTCAATTT</u> | <u>TCCACCATAT</u> | <u>TCATTTTATA</u> | <u>GCCATTATTT</u> |
| 141 | <u>AGGTGGCCCA</u> | <u>AAGTGAAATT</u> | <u>TACTCTTTTC</u>  | <u>TACTATTATA</u> | <u>AAAATTGAGG</u> | <u>ATGTTTGTGC</u> | <u>GGTACTTTGA</u> |
|     | <u>TCCACCGGGT</u> | <u>TTCACTTTAA</u> | <u>ATGAGAAAAG</u>  | <u>ATGATAATAT</u> | <u>TTTAACTCC</u>  | <u>TACAAAACAG</u> | <u>CCATGAAACT</u> |
| 211 | <u>TACGTCATTT</u> | <u>TTGTATGAAT</u> | <u>TGGTTTTTAA</u>  | <u>GTTTATTCGC</u> | <u>GATTTGGAAA</u> | <u>TGCATATCTG</u> | <u>TATTTGAGTC</u> |
|     | <u>ATGCAGTAAA</u> | <u>AACATACTTA</u> | <u>ACCAAAAATT</u>  | <u>CAAATAAGCG</u> | <u>CTAAACCTTT</u> | <u>ACGTATAGAC</u> | <u>ATAAACTCAG</u> |
| 281 | <u>GGTTTTTAAG</u> | <u>TTCGTTGCTT</u> | <u>TTGTAAATAC</u>  | <u>AGAGGGATTT</u> | <u>GTATAAGAAA</u> | <u>TATCTTTAAA</u> | <u>AAACCCATAT</u> |
|     | <u>CCAAAAATTC</u> | <u>AAGCAACGAA</u> | <u>AACATTTATG</u>  | <u>TCTCCCTAAA</u> | <u>CATATTCTTT</u> | <u>ATAGAAAATT</u> | <u>TTTGGGTATA</u> |
|     |                   |                   |                    |                   | <u>EcoRI</u>      |                   |                   |
|     |                   |                   |                    |                   | ~~~~~             |                   |                   |
| 351 | <u>GCTAATTTGA</u> | <u>CATAATTTTT</u> | <u>GAGAAAAATA</u>  | <u>TATATTCAGG</u> | <u>CGAATTCCAC</u> | <u>AATGAACAAT</u> | <u>AATAAGATTA</u> |
|     | <u>CGATTAAACT</u> | <u>GTATTAAAAA</u> | <u>CTCTTTTTAT</u>  | <u>ATATAAGTCC</u> | <u>GCTTAAGGTG</u> | <u>TTACTTGTTA</u> | <u>TTATTCTAAT</u> |
| 421 | <u>AAATAGCTTG</u> | <u>CCCCCGTTGC</u> | <u>AGCGATGGGT</u>  | <u>ATTTTTTCTA</u> | <u>GTAAAAATAA</u> | <u>AGATAAACTT</u> | <u>AGACTCAAAA</u> |
|     | <u>TTTATCGAAC</u> | <u>GGGGGCAACG</u> | <u>TCGCTACCCA</u>  | <u>TAAAAAAGAT</u> | <u>CATTTTATTT</u> | <u>TCTATTTGAA</u> | <u>TCTGAGTTTT</u> |
| 491 | <u>CATTTACAAA</u> | <u>AACAACCCCT</u> | <u>AAAGTCCTAA</u>  | <u>AGCCCAAAGT</u> | <u>GCTATGCACG</u> | <u>ATCCATAGCA</u> | <u>AGCCAGCC</u>   |
|     | <u>GTAAATGTTT</u> | <u>TTGTTGGGGA</u> | <u>TTTCAGGATT</u>  | <u>TCGGGTTTCA</u> | <u>CGATACGTGC</u> | <u>TAGGTATCGT</u> | <u>TCGGGTCGGG</u> |
| 561 | <u>AACCCAACCC</u> | <u>AACCCAACCC</u> | <u>ACCCAGTG</u>    | <u>AGCCAACTGG</u> | <u>CAAATAGTCT</u> | <u>CCACCCCGG</u>  | <u>CACTATCACC</u> |
|     | <u>TTGGGTTGGG</u> | <u>TTGGGTTGGG</u> | <u>TGGGGTCACG</u>  | <u>TCGGTTGACC</u> | <u>GTTTATCAGA</u> | <u>GGTGGGGGCC</u> | <u>GTGATAGTGG</u> |
| 631 | <u>GTGAGTTGTC</u> | <u>CGCACCACCG</u> | <u>CACGTCTCGC</u>  | <u>AGCCAAAAAA</u> | <u>AAAAAAGAA</u>  | <u>AGAAAAAAA</u>  | <u>GAAAAAGAAA</u> |
|     | <u>CACTCAACAG</u> | <u>GCGTGGTGGC</u> | <u>GTGCAGAGCG</u>  | <u>TCGGTTTTTT</u> | <u>TTTTTTTCTT</u> | <u>TCTTTTTTTT</u> | <u>CTTTTTCTTT</u> |
| 701 | <u>AACAGCAGGT</u> | <u>GGGTCCGGGT</u> | <u>CGTGGGGGCC</u>  | <u>GGAAAAGCGA</u> | <u>GGAGGATCGC</u> | <u>GAGCAGCGAC</u> | <u>GAGGCCCGGC</u> |
|     | <u>TTGTCGTCCA</u> | <u>CCCAGGCCCA</u> | <u>GCACCCCGG</u>   | <u>CCTTTTCGCT</u> | <u>CCTCCTAGCG</u> | <u>CTCGTCGCTG</u> | <u>CTCCGGGCCG</u> |

152 / 154

771 CCTCCCTCCG CTTCCAAAGA AACGCCCCC ATCGCCACTA TATACATACC CCCCCTCTC CTCCCATCCC  
GGAGGGAGGC GAAGGTTTCT TTGCGGGGGG TAGCGGTGAT ATATGTATGG GGGGGGAGAG GAGGGTAGGG

841 CCCAACCCCTA CCACCACCAC CACCACCACC TCCTCCCCC TCGCTGCCGG ACACGAGCT CCTCCCCCT  
GGGTTGGGAT GGTGGTGGTG GTGGTGGTGG AGGAGGGGGG AGCGACGGCC TGCTGCTCGA GGAGGGGGGA

911 CCCCCTCCGC CGCCGCCGGT AACCACCCCG CCCCTCTCCT CTTTCTTTCT CCGTTTTTTT TTTCGTCTCG  
GGGGGAGGCG GCGGCGGCCA TTGGTGGGGC GGGGAGAGGA GAAAGAAAGA GGCAAAAAA AAAGCAGAGC

981 GTCTCGATCT TTGGCCTTGG TAGTTTGGGT GGGCGAGAGC GGCTTCGTCTG CCCAGATCGG TGCGCGGGAG  
CAGAGCTAGA AACCGBAACC ATCAAACCCA CCCGCTCTCG CCGAAGCAGC GGGTCTAGCC ACGCGCCCTC

BamHI

~~~~~

1051 GGGCGGGATC TCGCGGCTGG CGTCTCCGGG CGTGAGTCGG CCCGGATCCT CGCGGGGAAT GGGGCTCTCG
CCCGCCCTAG AGCGCCGACC GCAGAGGCCC GCACTCAGCC GGGCCTAGGA GCGCCCCTTA CCCCAGAGC

BglII

~~~~~

1121 GATGTAGATC TTCTTTCTTT CTCTTTTTTG TGGTAGAATT TGAATCCCTC AGCATTGTTC ATCGGTAGTT  
CTACATCTAG AAGAAAGAAA GAAGAAAAAC ACCATCTTAA ACTTAGGGAG TCGTAACAAG TAGCCATCAA

HindIII NcoI

~~~~~

1191 TTTCTTTTCA TGATTTGTGA CAAATGCAGC CTCGTGCGGA GCTTTTTTGT **AGGTAGAAGC** TTACCATGG
AAAGAAAAGT ACTAAACACT GTTTACGTCG GAGCACGCCT CGAAAAAACA TC**CATCTTCG** AATGGTACC

KpnI-EcoRI - deletion underlined and restored NCO site in bold in vectors pJQ4.9,
pJQ3.2 and pJO6.3.

FIG. 50B

153 / 154

ALEURAIN\_deleted NPIR (Apoplast) Structure and Sequence



ALEURAIN-NPIR-DEL
 93 bp

+1 M A H A R V L L L A L A V L A T A A V A
 HindIII NcoI

~~~~~

1   AAGCTTACCA TGGCCACGC CCGCTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG  
 TTCGAATGGT ACCGGGTGCG GGCGCAGGAG GAGGACCGCG AGCGGCACGA CCGGTGCCGG CGGCAGCGGC

+1 V   A   S   S   R   A   A  
 NotI

~~~~~

71 TCGCCTCCTC CCGCGCGGCC GCC
 AGCGGAGGAG GGCGCGCCGG CGG

FIG. 51

154 / 154

SEE1 (Senescence enhanced) PROMOTER sequence

```

1   CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
      PstI                      XbaI      XbaI
      ~~~~~~
71  AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTGTG ATCTAGAACT CTAGATAGCA CAGCCACAGC
141 ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTG GAGACGGAGC TCTTTCCTAC CTCCTGACGT
211 TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC CCAACAAAAT ATCGTCCCCC
281 ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTCGCGCC GTTTTGTCT GAATCTCGCT TCCACTGGCC
      SmaI
      ~~~~~~
351 AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCG TCACCCCTGG CGTCATGGGA
421 TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAACCCAAT
      SphI
      ~~~~~~
491 TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTGTTG CTTAATTGAC TTTATTTTGG
561 TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT CGAGACGGAT AATAGGCTGG
631 CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGGCATTACA GCTGGAGGCT
      PstI
      ~~~~~~
701 TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG CGATGAGATG GGTATAAAAC
771 CCCCAGCACC GGGACGCGAG CTCCCGCCTA CCAGTACCAT CTCGCCTCGC TCCCCCTGCC GGACGACCCA
841 GTAAATACT GTTGCCCACT CGCCGGCGAG ATG
  
```

FIG.\_52

SEE1 (Senescence enhanced) PROMOTER plus vacuolar aleurain SIGNAL/NPIR sequence

```

1   CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
      PstI                      XbaI      XbaI
      ~~~~~~
71  AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTGTG ATCTAGAACT CTAGATAGCA CAGCCACAGC
141 ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTG GAGACGGAGC TCTTTCCTAC CTCCTGACGT
211 TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC CCAACAAAAT ATCGTCCCCC
281 ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTCGCGCC GTTTTGTCT GAATCTCGCT TCCACTGGCC
      SmaI
      ~~~~~~
351 AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCG TCACCCCTGG CGTCATGGGA
421 TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAACCCAAT
      SphI
      ~~~~~~
491 TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTGTTG CTTAATTGAC TTTATTTTGG
561 TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT CGAGACGGAT AATAGGCTGG
631 CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGGCATTACA GCTGGAGGCT
      PstI
      ~~~~~~
701 TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG CGATGAGATG GGTATAAAAC
771 CCCCAGCACC GGGACGCGAG CTCCCGCCTA CCAGTACCAT CTCGCCTCGC TCCCCCTGCC GGACGACCCA
      M A H G R I L F L A L A V L
841 GTAAATACT GTTGCCCACT CGCCGGCGAG ATGGCCACG GCCGCATCCT CTCTTGCGC CTCGCGGTCT
      BssHII
      ~~~~~~
      NotI
      ~~~~~~
      A T A A V A A A S L A D S N P I R P V T E R A
911 TGGCCACCGC CGCGGTGGCC GCCGCATCNT TGGCGGACTC CAACCCGATC CGGCCCCGTC CCGAGCGCGC
      NotI
      ~~~~~~
      A A
981 GGCCGCC
  
```

FIG.\_53